
NEWLY SEQUENCED NUCLEAR GENE (*XDH*) FOR INFERRING ANGIOSPERM PHYLOGENY¹

Cynthia M. Morton²

ABSTRACT

The nuclear gene xanthine dehydrogenase (*Xdh*) was sequenced for 247 genera representing all major lineages of angiosperms and “gymnosperms,” and the results were analyzed using likelihood and parsimony methods. The overall topology is mostly congruent with previously published trees based on chloroplast *rbcL*, *atpB*, and *matK* sequences. A total of 190 of the 253 nodes (71%) of the *Xdh* tree received bootstrap support greater than 50%. The likelihood tree was comparable in robustness to the *matK* topology, which exhibited 79% of the nodes with bootstrap support greater than 50%, and to the reported 7%–24% support observed for individual analyses of *rbcL*, *atpB*, and 18S ribosomal DNA clades. The number of parsimony-informative sites (1068, 69%) was similar to that of the *matK* (1083, 62%) study. The likelihood tree depicts angiosperms as monophyletic, with *Ceratophyllum* L. (Ceratophyllaceae) as sister to the rest of the flowering plants, followed successively by Amborellaceae, Nymphaeaceae, and Austrobaileyales clades as sisters to the remaining angiosperms. *Acorus* L. plus the remaining monocots, magnoliids, and Chloranthaceae diverge after the Austrobaileyales. Eudicots are supported and include a basal grade of Ranunculales–Proteaceae, Sabiaceae, Trochodendraceae, Buxaceae, Gunneraceae, and Dilleniaceae–Santalaceae, which are subsequent sister to the remaining eudicots. The remaining eudicots are split into two clades. The first clade consists of the Ericales, Cornales, and euasterids I and II (lamids–campanulids). The second clade consists of the following orders: Saxifragales, Myrtales–Caryophyllales–Cucurbitales, Crossosomatales, Geraniales, Rosales–Fabales–Fagales, Celastrales, Malpighiales, Brassicales–Malvales, Oxalidales, and Sapindales. *Xdh* data provided good support in the Caryophyllales, Ericales and Cornales, euasterids I (lamids), Magnoliales and Laurales, Malvales, Rutaceae, Oxalidales, Brassicales, and Sapindales. A future combined analysis of *Xdh* and other DNA data sets will have a strong potential to enhance resolution and internal support for angiosperm phylogenetics and provide insights into angiosperm evolution using biparental information.

Key words: Angiosperm, nuclear, phylogeny, systematics, xanthine dehydrogenase, *Xdh*.

DNA-based phylogenetic analyses of angiosperms have made use of a number of chloroplast markers over the past 15 years (Chase et al., 1993; Angiosperm Phylogeny Group, 1998; Soltis et al., 1999, 2000, 2003, 2005, 2007b; Savolainen et al., 2000; Angiosperm Phylogeny Group II, 2003; Hilu et al., 2003; Jansen et al., 2007). More recently, the basal angiosperms have been examined using chloroplast, nuclear, and mitochondrial markers (Qiu et al., 2006). Patterns have emerged regarding relationships among the angiosperms: (1) *Amborella* Baill. or *Amborella*–Nymphaeales represents the earliest diverging lineage of extant angiosperms, with Austrobaileyales the next clade to diverge; (2) *Ceratophyllum* L. is sister to either the monocots or eudicots; (3) magnoliids are composed of four orders (Canellales, Laurales, Magnoliales, and Piperales); (4) the monocots, a well-defined group; and (5) the Ranunculales, usually the first lineage in the eudicots, with Buxaceae, Caryophyllales, Cornales, Dilleniaceae, Ericales, euasterids, Gunnerales, Pro-

teaceae, rosids, Sabiaceae, Santalales, and Trochodendraceae following in various arrangements and forming well-supported clades (Angiosperm Phylogeny Group II, 2003; Hilu et al., 2003; Soltis et al., 2005). Not all lineages are stable in their positions on individual or combined gene trees, and many have weak support. Major clades within basal angiosperms that have varied in position include the following: *Amborella*, Austrobaileyales, *Ceratophyllum*, magnoliids, monocots, and Nymphaeaceae (Angiosperm Phylogeny Group II, 2003; Hilu et al., 2003; Soltis et al., 2003; Qiu et al., 2006; Moore et al., 2007). Moreover, the relationships among the major clades of eudicots (Caryophyllales, Cornales, Dilleniaceae, Ericales, euasterids, Proteaceae, rosids, Sabiaceae, and Santalales) remain uncertain (Angiosperm Phylogeny Group II, 2003; Hilu et al., 2003; Soltis et al., 2003; Qiu et al., 2006). Resolving these relationships is essential for systematics and for tracing character evolution. Furthermore, many other fields of study rely on a well-supported angiosperm phylogeny, so it

¹ We thank Abigail Mazie, Jodie Rabbitt, Sarah Suchy, Sandra Newman, and Hyosig Won for providing technical support. Plant material was generously provided by Phipps Conservatory and Botanical Gardens, Fairchild Tropical Botanic Garden, University of California Botanical Garden at Berkeley, and Dr. Yin-Long Qiu. Funding for this project was provided in part by the Heinz Endowment.

² Head of Section of Botany, Carnegie Museum of Natural History, 4400 Forbes Avenue, Pittsburgh, Pennsylvania 15213-4080, U.S.A. mortoncm@verizon.net.

doi: 10.3417/2008107

is crucial that phylogenetics produce a well-resolved topology using molecular and nonmolecular features.

Studies of broad angiosperm phylogenetics have centered on using chloroplast genes individually and in combination (Chase et al., 1993; Soltis et al., 1997, 2000; Savolainen et al., 2000; Hilu et al., 2003; Jansen et al., 2007). More recently, mitochondrial genes have been incorporated into basal angiosperm phylogenies (*rpoC2*, Qiu et al., 2006). Nuclear genes have hardly been used, and only the 18S gene has been extensively incorporated into multigene analyses investigating angiosperm phylogeny (Soltis et al., 1997). On its own, 18S has provided very limited resolution and low internal support for many major clades (Soltis et al., 1997). Studies of *rbcL* and *atpB* chloroplast genes have provided phylogenies with more resolution than individual genes; however, topological incongruence does occur between the genes (Savolainen et al., 2000). Analysis of the chloroplast gene *matK* provided a tree with greater resolution and more support for major clades, but topological incongruences were observed among *matK*, *rbcL*, and *atpB*. Combining chloroplast genes in multigene data sets for analysis has improved both resolution and internal support (Savolainen et al., 2000; Burleigh et al., 2009). Hilu et al. (2003) compared bootstrap/jackknife values for major nodes within the angiosperm trees and found the following supported percentages for these genes: *rbcL* (24%), *atpB* (14%), 18S (7%), *rbcL/atpB* (55%), and *rbcL/atpB/18S* (83%). When using chloroplast genes, one must be a bit cautious because these genes are maternally inherited and, due to issues such as hybridization, the resulting phylogenetic reconstructions may not reflect the “true phylogeny” or correct topology. The first angiosperm consensus tree based on combined chloroplast (*rbcL* and *atpB*) and nuclear (18S) genes was conducted by Soltis et al. (2000). Although 18S is a very slowly evolving gene, the combined chloroplast and nuclear consensus tree may be considered the currently most reliable overall angiosperm phylogeny because it is based on both maternally and biparentally inherited genes. However, many of the interrelationships in the angiosperm and land plant phylogenies are still controversial, and it is imperative to use information from different biparentally inherited genes to clarify and enhance the results derived largely from chloroplast genes.

Xanthine dehydrogenase is a member of the molybdenum hydroxylase family of enzymes, which are thought to play important metabolic roles in purine metabolism and hormone biosynthesis. Molybdenum (Mo) is essential for (nearly) all organisms and functions in more than 40 enzymes by catalyzing

diverse redox reactions. Only four of those enzymes have been found in plants: nitrate reductase; aldehyde oxidase(s); xanthine dehydrogenase, which is involved in purine catabolism and stress reactions; and sulphite oxidase (Mendel & Hänsch, 2002). Among Mo enzymes, the alignment of amino acid sequences permits the definition of domains that are well conserved, although they are variable. This study amplified approximately 1265 bp of the xanthine dehydrogenase gene (*Xdh*), which is located on chromosome 4, contig fragment no. 82, and is approximately 4083 bp long (information obtained by comparing published *Arabidopsis* and this study's sequences). The sequenced segment lies from 358 to 1623 bp including intron 1–3 and exon 2–4. The *Arabidopsis* (DC.) Heynh. sequence contains 13 introns and 14 exons. Hesberg et al. (2004) found that *Arabidopsis* has a 46%–47% identity to human and bovine *Xdh* and therefore is homologous to *Xdh* from other organisms. This examination segment contains approximately 11 indel regions. To date, only one copy of the gene has been reported, except in silkworms and *Arabidopsis*, in which a second copy has been found. In *Arabidopsis*, Hesberg et al. (2004) reported that the physicochemical properties of AtXDH1 and AtXDH2 might be nearly identical because of their high degree of sequence similarity (94%).

Markers from the nuclear genome that can be used on a broad taxonomic scale without problems of paralogy are limited in number. The possibility of paralogy exists for all genes, not just nuclear ones, because gain and loss of genes by various processes are continuous and all genes have some probability of being duplicated or lost. One consequence of paralogy is that gene trees derived from samples of single genes from each species may not match the species tree. Having more than one gene tree to compare the angiosperm phylogeny is a good indicator of having paralogous genes. We have used this test of congruence among genes to evaluate duplications or losses within *Xdh*.

This analysis provides the first angiosperm plant gene tree using *Xdh*. More recently other researchers have been exploring the utility of *Xdh* for “gymnosperm” phylogenetics and have found it to be similar to inferences using other genes and to be highly informative (Peery et al., 2008).

We compared the topology obtained from this gene to previously published topologies from *rbcL*, *atpB*, and *matK* genes. This assessment was done to achieve two major goals: (1) compare maternally inherited genes to a biparental gene to see if the topologies are more, equally, or less resolved,

supported, and congruent, and (2) examine and compare the topologies for basal angiosperms and eudicots among the data sets for a greater understanding of seed plant relationships.

MATERIALS AND METHODS

TAXON SAMPLING

Several different publications were consulted for taxon sampling schemes, including Savolainen et al. (2000), Soltis et al. (2000), Hilu et al. (2003), and Qiu et al. (2006). In total, 253 species representing 164 families and 247 genera of angiosperms and gymnosperms were sampled, and all but one was sequenced specifically for this study. See Appendix 1 for a list of taxa with accession data and GenBank numbers.

DNA ISOLATION, GENE AMPLIFICATION, AND SEQUENCING

For most samples, total DNA was isolated from fresh or herbarium specimens, although some was acquired from previously extracted samples. Leaves were ground with a mortar and pestle and subsequently extracted using the DNEasy plant DNA extraction kit (Qiagen, Valencia, California, U.S.A.) following the manufacturer’s protocol. Specific polymerase chain reaction (PCR) and sequencing primers were initially designed using the *Arabidopsis* plant sequences and other organisms available in GenBank and used by Rodríguez-Trelles et al. (2003). Additional primers were redesigned as needed. See Table 1 for a list of primers used for amplification and sequencing. Cloning was performed on selected taxa using the pGem T-easy vector system II kit (Promega Corporation, Madison, Wisconsin, U.S.A.) following the manufacturer’s protocol. PCR was performed with 10 µmol/L primers in 25-µL reactions. In general, it was conducted using the following program: 96°C for 1 min. followed by 34 cycles of 94°C for 30–60 sec., 48°–60°C for 1 min., 68°C for 60–80 sec., and finally 72°C for 7 min. For some taxa, the initial 96°C for 1 min. and the final 72°C for 7 min. were eliminated. Primers used for amplification of clones were M13F and M13R. For sequencing reactions, the Big Dye Terminator Sequencing 3.1 kit (Applied Biosystems, Foster City, California, U.S.A.) was used, and fragments were separated on an ABI3730 DNA Analyzer from Applied Biosystems.

PHYLOGENETIC ANALYSIS

The *Xdh* gene was analyzed using both maximum parsimony (MP) and maximum likelihood (ML). For

Table 1. Primers used for the amplification along with their sequences. The *Xdh* gene was split into two halves for amplification (forward and reverse). All taxa used in this study were amplified with a combination of the primers listed below, except for primer 1299-GF, which was designed to amplify the back half in gymnosperm taxa.

Primer name	Sequence
*471F	5'-CAATGTGGRTTTCTYACYCCYGG-3'
*486F	5'-ACYCCYGGKTTTTRTIATGTCIATGTA-3'
497F	5'-TTGTGATGTCCGATGTATGC-3'
975R	5'-TGCTCCWGCAGCCATCCAGAG-3'
*1362R	5'-TCWGATATTGGACTAGCWGT-3'
*1365R	5'-TTYAARTCHGATATYGGACTRGC-3'
*1296F	5'-AARTGGTTYGCGGIACACARAT-3'
1299F	5'-TGGTTTGCTGGGACACAAAT-3'
1299-1F	5'-TGGTTTGCHGGSAMHCARAT-3'
1299-GF	5'-TGGTTTGCTGGVAVYCARAT-3'
*1869R	5'-CGAAAYTCMACCATTCMCCAGG-3'
1872R	5'-CGAAAYTCMACCATTCMCC-3'
1869-1R	5'-CKAAWYTCMACCATTCMCCAGG-3'
1354R	5'-CAGAAACTATCCATKICCC-3'

* Amplification and sequencing primers used in more than 75% of the taxa examined.

the MP analysis, a heuristic search with PAUP* (Swofford, 2004) was completed using the following strategy: addition sequence random, number of replicates 100, tree bisection-reconnection branch swapping, MULTREE option on, and MAXTREE auto-increased by 100.

For likelihood analyses, the program MODEL-TEST v. 3.6 (Posada & Crandall, 1998) was used to select the models of nucleotide evolution. The program GARLI v. 0.96 (Genetic Algorithm for Rapid Likelihood Inference; Zwickl, 2006) implementing the general time reversible (GTR) substitution model, with the base frequency and proportion of invariable sites estimated, was used for the *Xdh* gene data set.

The parsimony bootstrapping analyses (100 replicates) were conducted in PAUP* (Swofford, 2004) using simple taxon addition, one tree held at each step during stepwise addition, tree bisection-reconnection branch swapping, MULTREE option on, and no upper limit of MAXTREE set.

RESULTS

CLONES

Among plants, only *Arabidopsis* has been previously reported to have two copies of *Xdh*, and these copies have a high degree of similarity (94%). Hesberg et al. (2004) reported that the physicochemical properties of *Arabidopsis* AtXDH1 and AtXDH2

might be nearly identical because of their high degree of similarity (94%). For this study, over 80 clones were sequenced representing 20 taxa. Taxa that had more than one copy had 91%–99.8% similarity among sequences, such as *Coriaria sarmentosa* G. Forst. (sp.) of the Cucurbitaceae having a 95% similarity among cloned sequences. Only *Tacca* J. R. Forst. & G. Forst. and *Polygala* L. had similarities of 71% and 58%, respectively, among the cloned sequences. In both taxa, once the highly divergent sequence was removed, homology among the remaining sequences was 98% and 99%, respectively.

COMPARING METHODS OF ANALYSIS

Sequence variability

Approximately 1265 bp were sequenced, resulting in 1557 aligned characters due to the insertion of gaps. Sequences were aligned manually using Sequencher 4.7 (Gene Codes Corporation, Ann Arbor, Michigan, U.S.A). Eleven indels ranging from 3 to 66 bp in length were inserted in the alignment. Of the aligned characters, 1187 (76%) were variable, 1068 (69%) were potentially parsimony-informative, and approximately 180 (12%) were missing. The transition:transversion ratio in *Xdh* was 1.10, indicating that a high level of saturation was not reached (ratio of 0.4 and below is an indication of highly saturated sequences, according to Holmquist [1983]).

Phylogenetic results

Parsimony trees containing 253 taxa produced 21,585 trees of 31,337 steps in length, with a consistency index (CI) = 0.11 and a retention index (RI) = 0.50 (940 hours used). Homoplasy levels are comparable with those from Soltis et al.'s (2000) three-gene data set (CI = 0.12, 567 taxa) and Hilu et al.'s (2003) *matK* gene data set (CI = 0.14, 374 taxa).

The discussion is based on the likelihood tree because of its better resolution and support, and the results from the MP topology will be contrasted only when relevant. Bootstrap support of 50%–74% is considered low, 75%–84% moderate, and greater than 85% high (Chase et al., 2000).

The likelihood tree (Figs. 1–8) has a total of 190 of the 253 nodes (71%) that received bootstrap support of 50% or greater, and support levels averaged 71%. The likelihood tree depicts the angiosperms as monophyletic (80%) with *Ceratophyllum* (Ceratophyllaceae) as a sister to the rest of the flowering plants, followed successively by *Amborella* (Amborellaceae), a clade incorporating *Brasenia* Schreb., *Nuphar* Sm., and *Nymphaea* L. (Nymphaeaceae) and an *Austro-*

baileya C. T. White–*Kadsura* Juss.–*Schisandra* Michx. (Austrobaileyales) clade as sister to the remaining angiosperms. *Acorus* L. (73%) is sister to the remaining monocots (74%). A weakly supported magnoliid clade (42%) consists of Piperales and Canellales (56%) and Laurales and Magnoliales (90%). *Acorus* plus the remaining monocots, magnoliids, and Chloranthaceae sequentially diverge after the Austrobaileyales. Eudicots include a basal grade of Ranunculales and Proteaceae, Sabiaceae, Trochodendraceae, Buxaceae, Gunneraceae, and Dilleniaceae and Santalaceae, which form a grade to the remaining eudicots. The grade of taxa containing Ranunculales and Proteaceae, Trochodendraceae, and Gunneraceae to the eudicots received 72%, 71%, and 72% support, respectively. The remaining eudicots are split into two clades. The first clade consists of the Ericales, Cornales, and euasterids I and II (lamids and campanulids), i.e., the asterid clade (Asteridae, Fig. 5). The second clade consists of the following orders: Saxifragales, Myrtales–Caryophyllales–Cucurbitales, Crossosomatales, Geraniales, Rosales–Fabales–Fagales, Celastrales, Malpighiales, Brassicales–Malvales, Oxalidales, and Sapindales (Figs. 6–8).

DISCUSSION

EARLY-DIVERGING ANGIOSPERMS

Early-diverging angiosperms have been the subject of intensive molecular studies (Chase et al., 1993; Soltis et al., 1997, 1999, 2000, 2005; Hoot et al., 1999; Mathews & Donoghue, 1999, 2000; Parkinson et al., 1999; Qiu et al., 1999, 2000, 2005, 2006; Barkman et al., 2000; Graham & Olmstead, 2000; Savolainen et al., 2000; Zanis et al., 2002, 2003; Borsch et al., 2003; Goremykin et al., 2003; Hilu et al., 2003; Aoki et al., 2004; Stefanović et al., 2004; Löhne & Borsch, 2005). Recent studies have found the earliest diverging lineages of the angiosperms to be members of the Amborellaceae and Nymphaeaceae, either as a clade or grade, followed by the Austrobaileyales (Savolainen et al., 2000, *atpB*; Soltis et al., 2000, 2005; Hilu et al., 2003; Qiu et al., 2006). The *Xdh* gene had a topology supported by 81% bootstrap values for the monophyly of the angiosperms consisting of *Ceratophyllum*, followed by *Amborella*, Nymphaeaceae, and Austrobaileyales (Fig 1). The position of *Ceratophyllum* has varied widely in previous studies: sister to either all angiosperms (Chase et al., 1993; Savolainen et al., 2000 [*rbcL*, *atpB/rbcL*]), sister to monocots (Zanis et al., 2002, 2003; Qiu et al., 2005; Soltis et al., 2005), sister to Chloranthaceae (Qiu et al., 2006), sister to eudicots

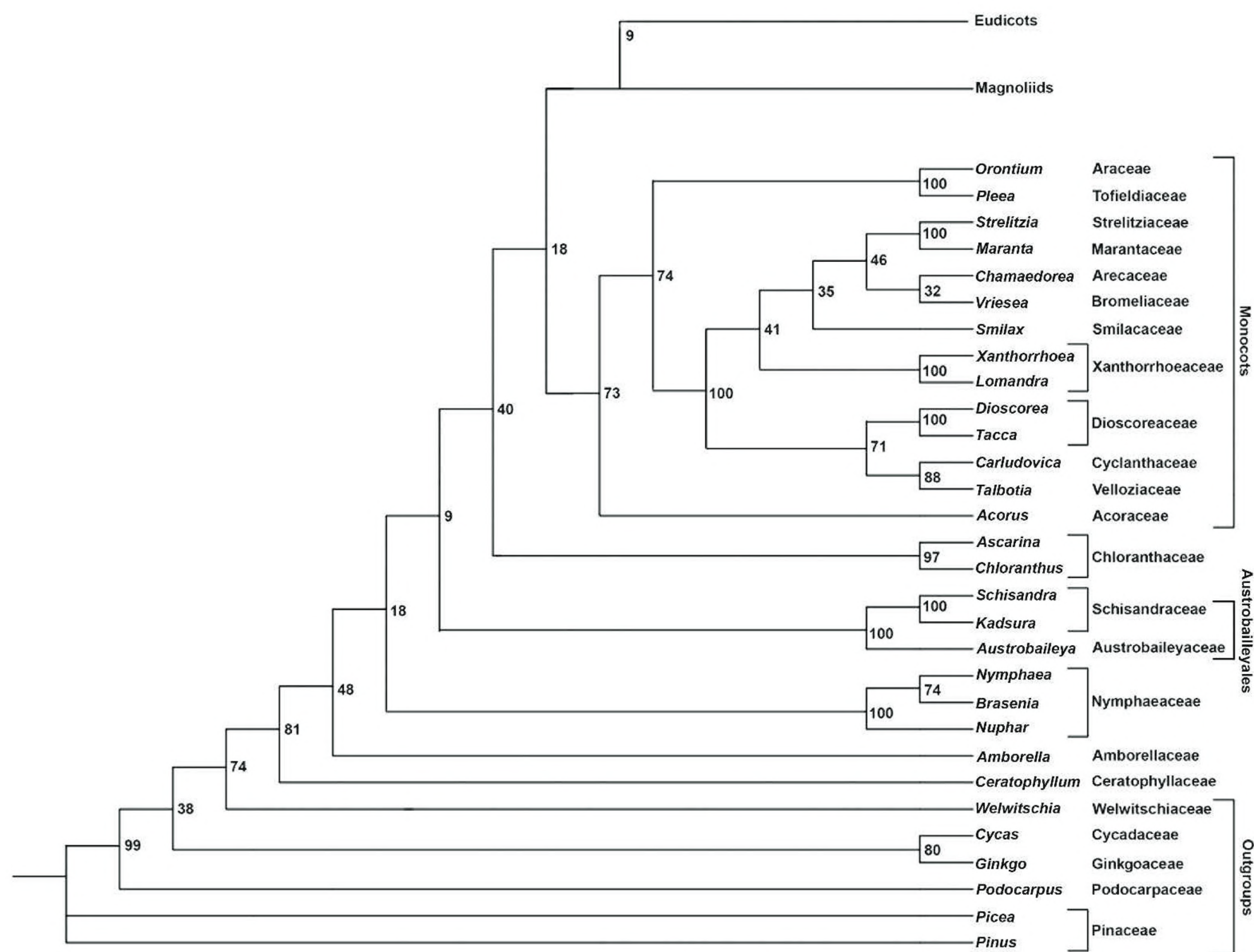


Figure 1. Bootstrap consensus tree of likelihood analyses among basal angiosperm lineages. The numbers above the branches are likelihood bootstrap values.

(Soltis et al., 2000; Hilu et al., 2003; Qiu et al., 2005, 2006), or sister to *Acorus*, which is sister to the remaining monocots (Savolainen et al., 2000 [*atpB*]). Only one sample of *Ceratophyllum* was used in our study, which may influence the placement of this taxon in this analysis. The placement of *Amborella* and Nymphaeales followed by Austrobaileyales is in agreement with most cited studies. In one of the most intensive studies of the earliest diverging lineages, using three mitochondrial, one nuclear, and four chloroplast genes (Qiu et al., 2006), different inferences among gene data sets were shown. Comparison of the mitochondrial and chloroplast data sets found that mitochondrial genes supported *Amborella*–Nymphaeales, while chloroplast genes supported *Amborella* as sister to remaining angiosperms. Depending on the data set, *Ceratophyllum* was either sister to Chloranthaceae, with a likelihood bootstrap value of 26%, or sister to eudicots, with a likelihood bootstrap value ranging from 38% to 51%. Only one nuclear gene was examined in the study from Qiu et al. (2006), and comparisons were not made with the mitochondrial or chloroplast topologies. Previous studies, along with the present

research, suggest that resolving these relationships will require more research.

CHLORANTHACEAE

The position of Chloranthaceae has varied in previous studies, from sister to all of the euangiosperms (Qiu et al., 2006; Cantino et al., 2007), sister to *Ceratophyllum* (Qiu et al., 2006), sister to monocots (Hilu et al., 2003), sister to magnoliids and eudicots (Zanis et al., 2002, 2003; Soltis et al., 2005), sister to magnoliids (Savolainen et al., 2000 [*atpB* data set]), or sister to monocots and magnoliids (Savolainen et al., 2000 [using the *rbcL* or *atpB/rbcL* data sets]; Soltis et al., 2000). Most of these placements have had very low to only moderate bootstrap support. This family exhibits numerous distinctive morphological characters and putative synapomorphies, and its fossil record can be traced back to the Early Cretaceous (Friis et al., 1986; Pedersen et al., 1991; Doyle et al., 2003). The *Xdh* likelihood tree found that Chloranthaceae, represented here by two genera, forms a strongly supported clade (97%) that is sister to all monocots, magnoliids, and eudicots (Fig. 1). The most parsimonious tree

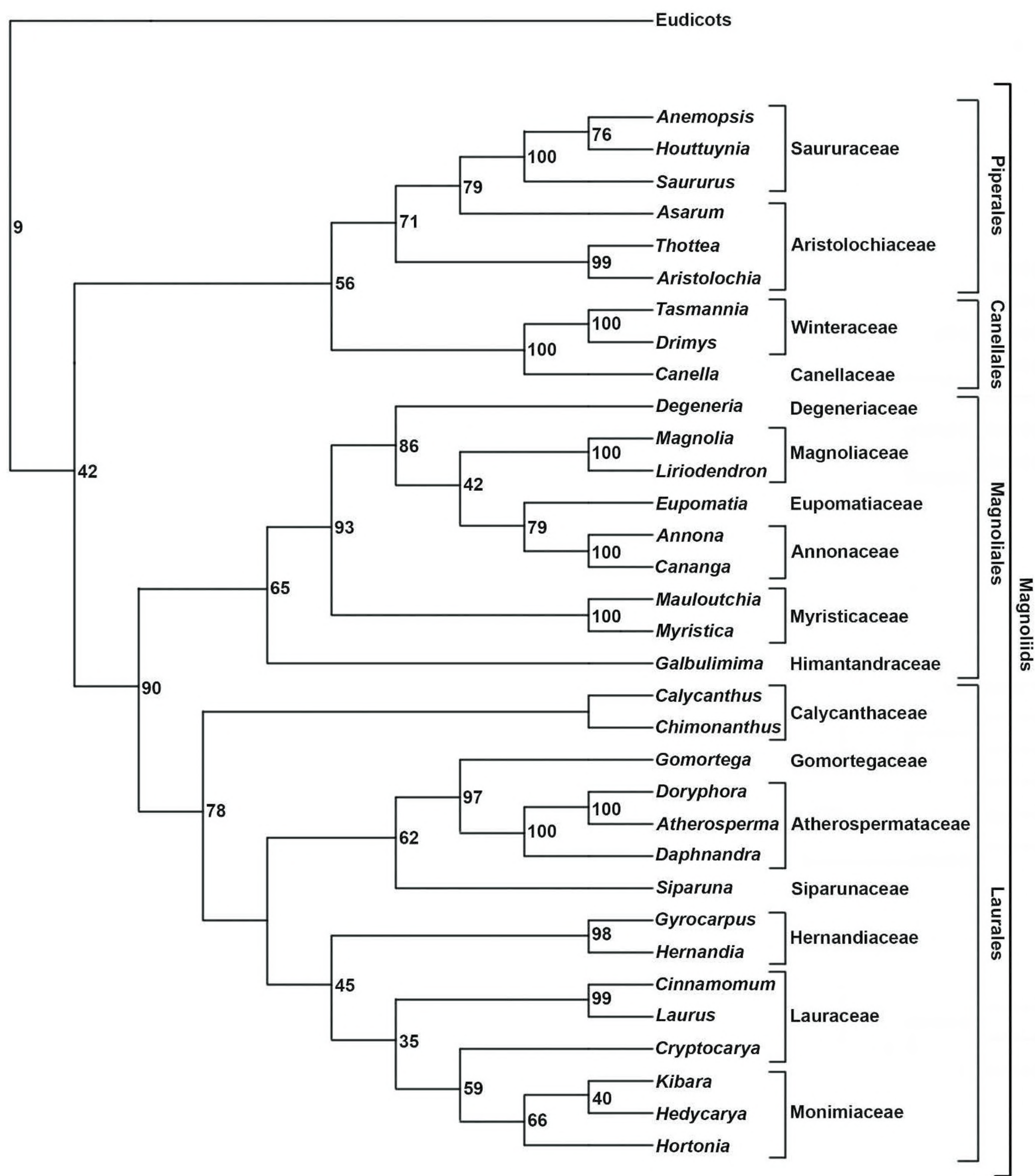


Figure 2. Bootstrap consensus tree of likelihood analyses among magnoliids. The numbers above the branches are likelihood bootstrap values.

(MPT) differs in that Chloranthaceae was sister to all eudicots.

ACORUS PLUS REMAINING MONOCOTS

The *Xdh* sequences provided moderate support (73%) for the monophyly of the monocots, with Alismatales (Araceae and Tofieldiaceae) followed by *Acorus* as sister to the remaining monocots (Fig. 1). This topology has been found in most single and multigene analyses (Chase et al., 2000, 2006; Savolainen et al., 2000 [*rbcL*, *atpB/rbcL*]; Soltis et al., 2000, 2005; Hilu et al., 2003). The monophyly of

the Alismatales received high support (100%), along with the monophyly of the remaining monocots (100%). The internal structure of the monocots is represented by two clades. The first clade contains Dioscoreales (100%) and Pandanales (Cyclanthaceae and Velloziaceae, 88%) with 71% support. The second clade includes sister to Asparagales (Xanthorrhoeaceae, 100%), followed by Liliales (Smilacaceae), and two nested sister clades, one consisting of Poales (Bromeliaceae) and Arecales with weak support (32%) and the other of Zingiberales (Marantaceae and Strelitziaceae, 100%). These relationships are highly congruent with those suggested by Hilu et al. (2003).

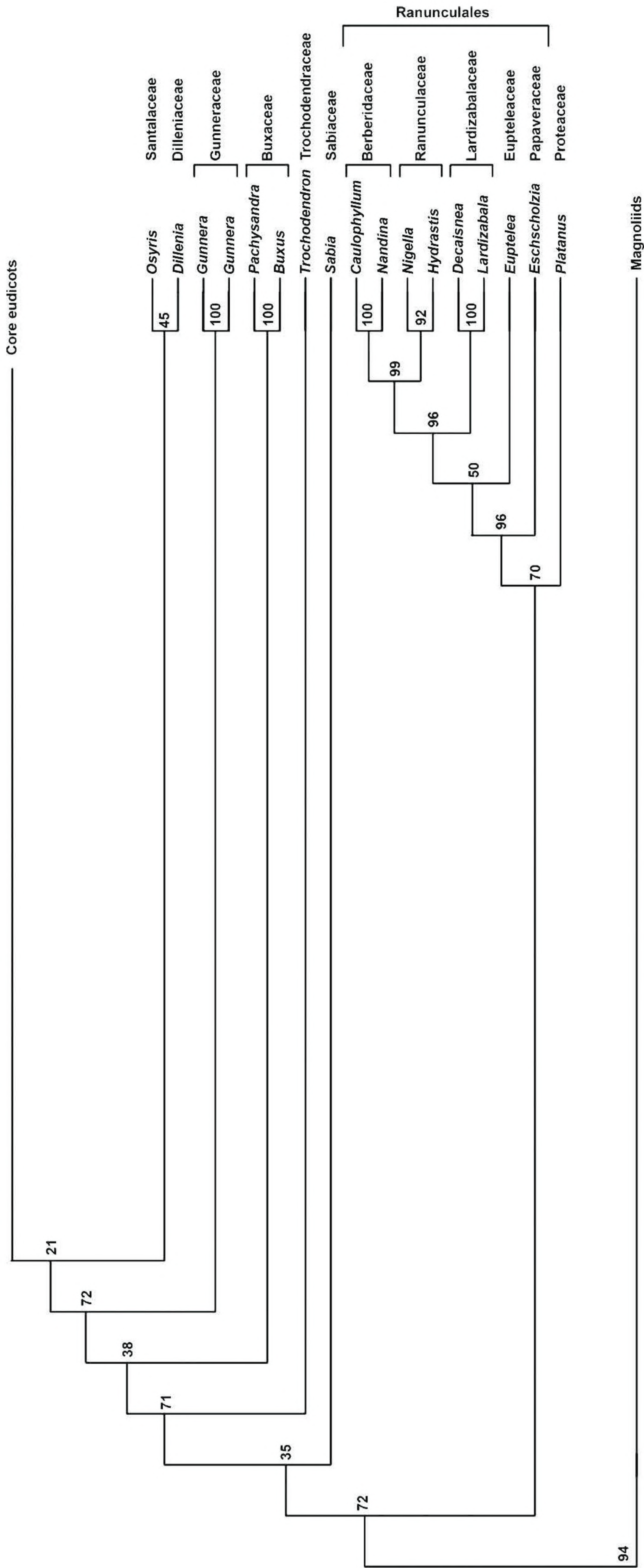


Figure 3. Bootstrap consensus tree of likelihood analyses among basal grade of eudicots. The numbers above the branches are likelihood bootstrap values.

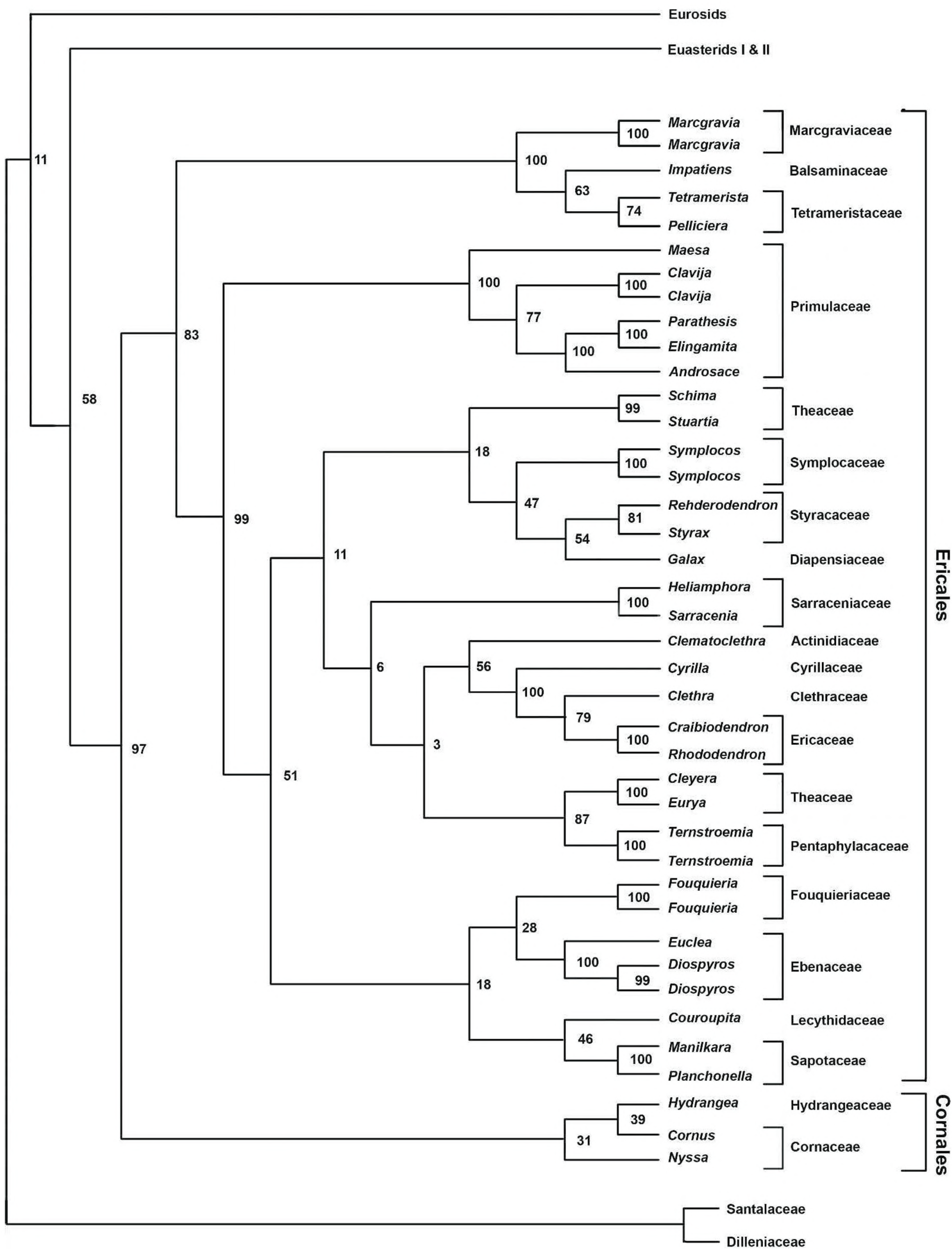


Figure 4. Bootstrap consensus tree of likelihood analyses among Ericales. The numbers above the branches are likelihood bootstrap values.

Our results are also consistent with those of Chase et al. (2000, 2006) based on seven genes, 18S rDNA, partial 26S, *rbcL*, *atpB*, *matK*, *ndhF*, and *atpI*. The exact positions of some orders are slightly different; for example, in Chase et al. (2006), Dioscoreales and

Pandanales are sister to Liliales, while in trees based on *Xdh*, *Smilax* L. (Liliales) is sister to the Poales and Arecales. The monocots are sister to the magnoliids (a clade consisting of Piperales, Canellales, Laurales, and Magnoliales).

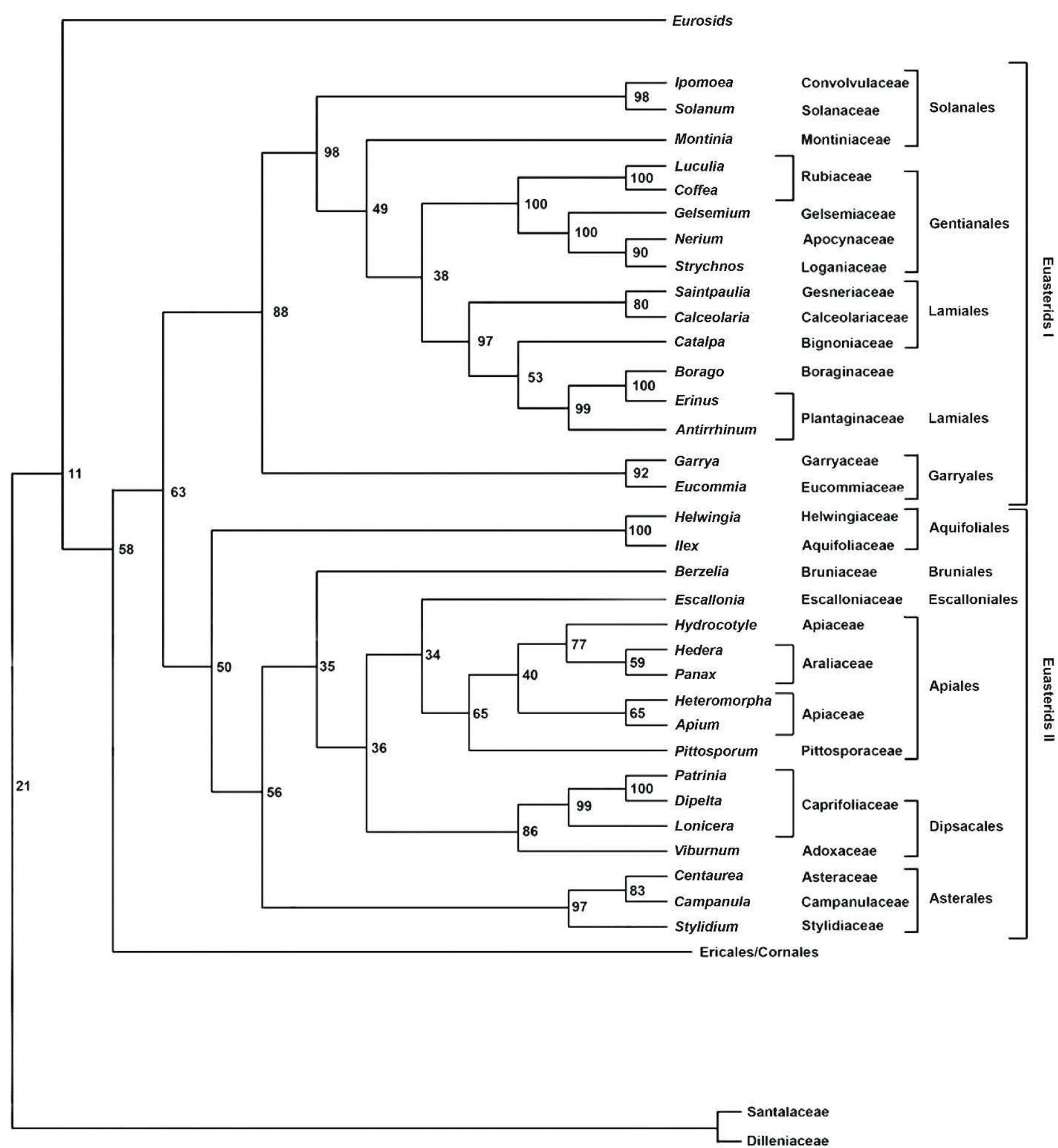


Figure 5. Bootstrap consensus tree of likelihood analyses among euasterid I and II. The numbers above the branches are likelihood bootstrap values.

MAGNOLIIDS

Recent studies provide strong evidence for a clade including only Piperales, Canellales, Laurales, and Magnoliales (generally referred to as the magnoliids or eumagnoliids). The *Xdh* topology is in agreement with this circumscription, although some earlier phylogenetic studies (Chase et al., 1993; Savolainen et al., 2000 [*rbcL*, *atpB/rbcL*]; Soltis et al., 2000) did not recognize the monophyly of the magnoliids or the relationships among the four orders (Fig. 2). Qiu et al. (1999, 2000, 2006) identified the sister relationships between Magnoliales and Laurales and between Canellales and Piperales, as well as the monophyly of all of the magnoliids. Hilu et al. (2003) also recovered the monophyly of the magnoliids, with Magnoliales–Laurales and Piperales–Canellales as sister groups, but with low levels of support. The *Xdh*

likelihood tree has high levels of support for the Magnoliales–Laurales sister-group relationship (90%), whereas the Canellales–Piperales sister-group relationship received lower support (56%). These sister-group relationships have been found in several other analyses (Mathews & Donoghue, 1999, 2000; Qiu et al., 1999, 2000, 2005, 2006; Barkman et al., 2000; Graham & Olmstead, 2000; Zanis et al., 2002, 2003; Borsch et al., 2003, 2005; Hilu et al., 2003; Löhne & Borsch, 2005; Soltis et al., 2005; Graham et al., 2006).

In the clade consisting of the Canellales, Winteraceae formed a strongly supported clade (100%) with *Canella* P. Browne (100%). The Piperales (71%) contained members of Saururaceae and Aristolochiaceae. The position of *Asarum* L. (Aristolochiaceae) as sister to the Saururaceae (100%)

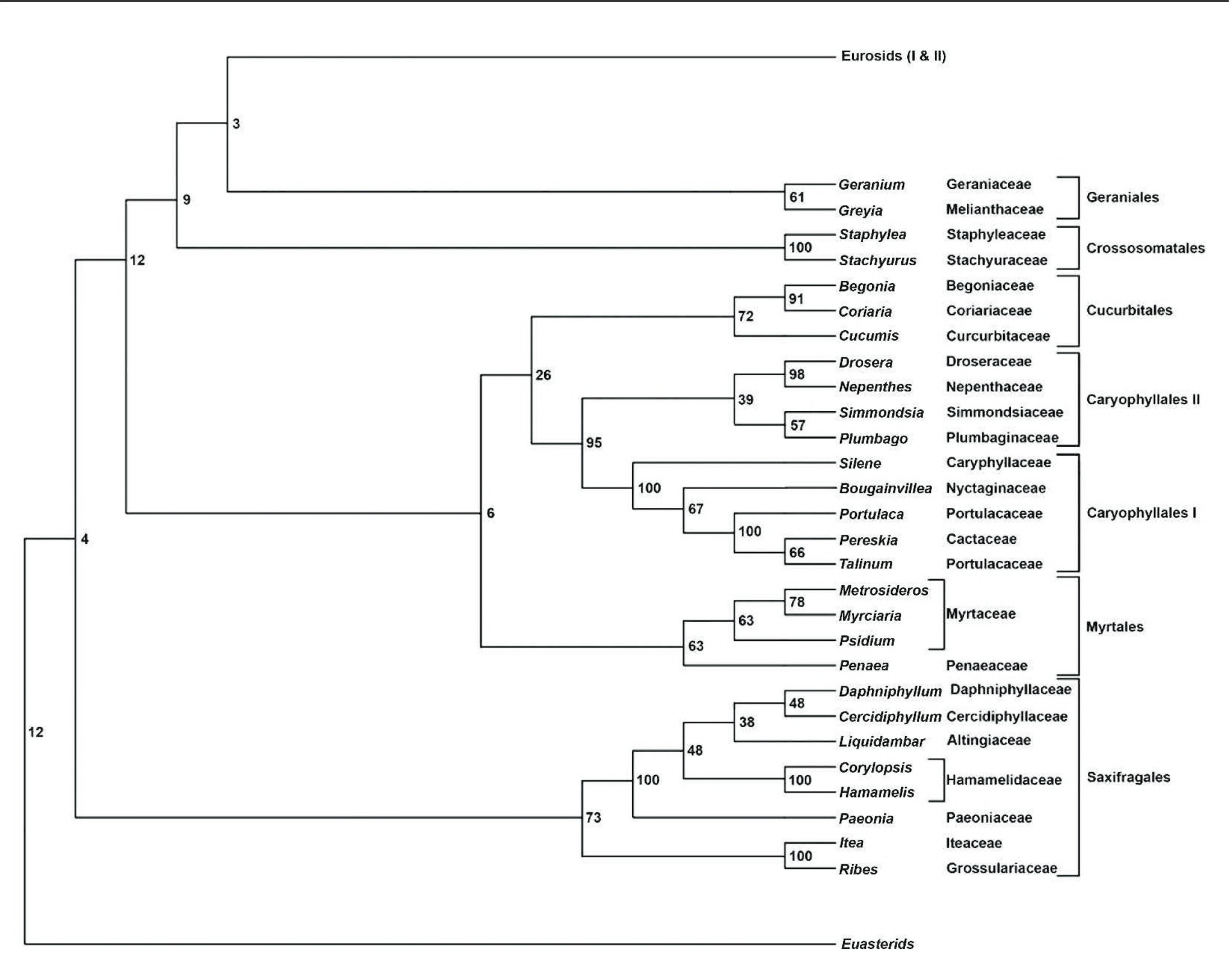


Figure 6. Bootstrap consensus tree of likelihood analyses among the core eudicots. The numbers above the branches are likelihood bootstrap values.

was different than that in most other analyses. However, this position for *Asarum* was also found in the study by Hilu et al. (2003), but, with increased taxon sampling, Qiu et al. (2006) showed that *Asarum* regrouped with members of Aristolochiaceae. Morphological characters clearly place *Asarum* in Aristolochiaceae, with aristolochic acids, sepals connate, filaments slightly adnate to the style, ovary inferior, and ovules numerous among their common traits (Judd et al., 2007). Limited taxon sampling may be influencing the placement of *Asarum* in this analysis. Increasing the sampling of Aristolochiaceae, Piperaceae, and Saururaceae may result in different inferences.

The Magnoliales formed a weakly supported clade (65%), and all families with more than one genus are monophyletic. The Annonaceae, Degeneriaceae, Eupomatiaceae, Himantandraceae, Magnoliaceae, and Myristicaceae have occupied various positions in most major analyses (Qiu et al., 1999, 2000, 2005, 2006; Zanis et al., 2002; Hilu et al., 2003). *Xdh* sequences provide evidence that Himantandraceae is sister to the remaining members of the order, followed by Myristicaceae, Degeneriaceae, and Magnoliaceae,

which are sister to the clade Eupomatiaceae and Annonaceae.

The Laurales formed a moderately supported clade (78%) with all multitaxa families remaining monophyletic, except for the Lauraceae. The nonmonophyly of Lauraceae was also found by Qiu et al. (2006) using several taxa, but Chanderbali et al. (2001) examined additional taxa and found a clade containing *Cryptocarya* R. Br. to be basal in the Lauraceae. The placements of Atherospermataceae, Calycanthaceae, Gomortegaceae, Hernandiaceae, Lauraceae, Monimiaceae, and Siparunaceae of the Laurales have varied in most major analyses, and their relationships will remain unresolved until more examinations are completed (Qiu et al., 1999, 2000, 2005, 2006; Chanderbali et al., 2001; Zanis et al., 2002; Hilu et al., 2003).

Although the magnoliid grouping has been found in many other molecular analyses, sound morphological or phytochemical characters to define the groups are lacking (Doyle & Endress, 2000). Soltis et al. (2005) cited some chemical compounds such as neolignan and P-type sieve-tube plastids for most of the families, and Stevens (2001, onward) mentions

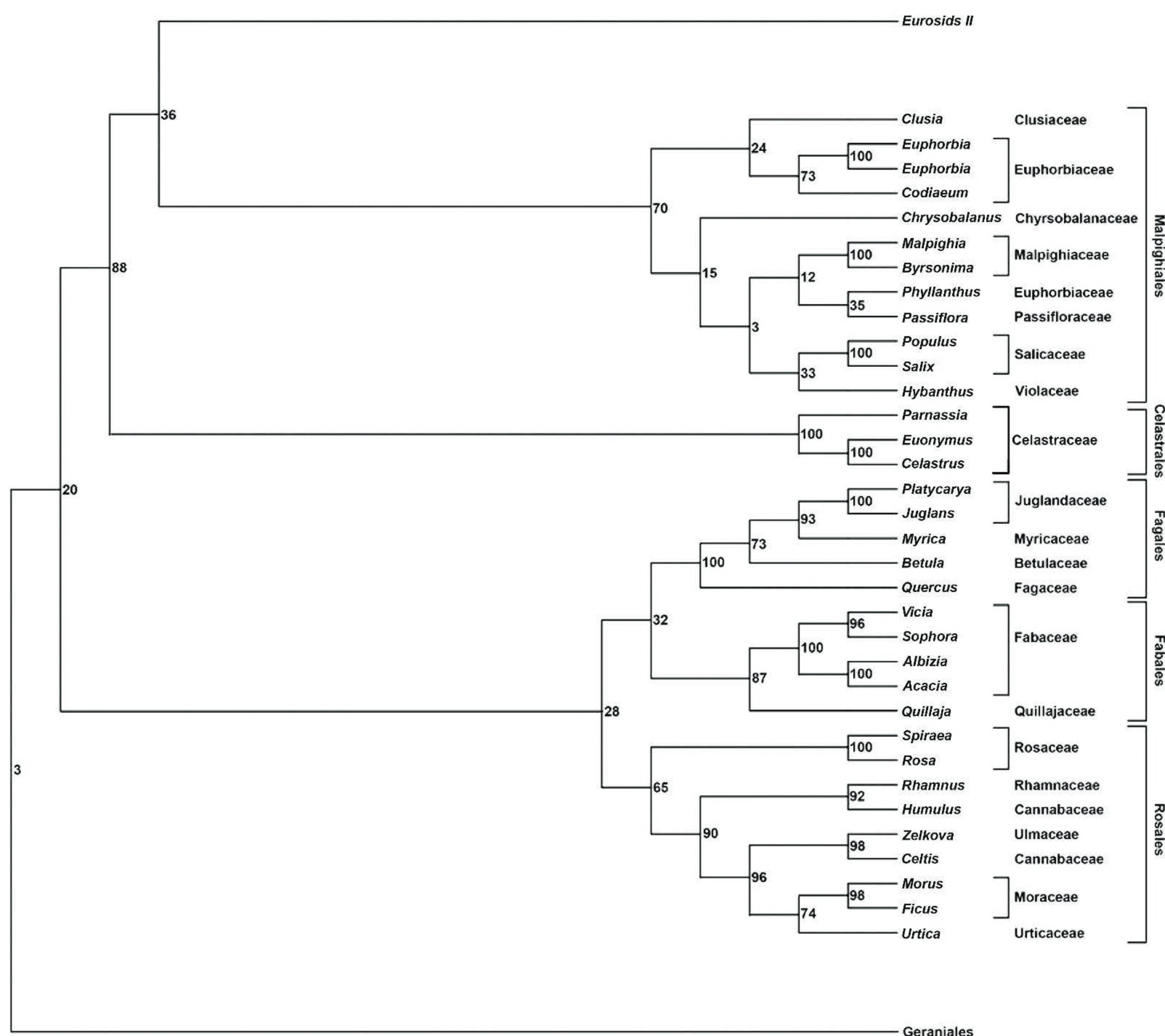


Figure 7. Bootstrap consensus tree of likelihood analyses among eurosid I. The numbers above the branches are likelihood bootstrap values.

chemical compounds such as neolignan and embryological features, but there are no unambiguous synapomorphies for this clade. The *Xdh* analysis has provided additional evidence for the monophyly and relationships among the magnoliids.

EUDICOTS

Early-diverging eudicots

The early-diverging eudicot taxa—Ranunculales, Proteaceae, Sabiaceae, Trochodendraceae, Buxaceae, and Gunneraceae—have been consistently recovered together as mainly a grade in a number of analyses (Chase et al., 1993; Hoot et al., 1999; Savolainen et al., 2000 [*atpB*, *atpB/rbcL*]; Soltis et al., 2000, 2003, 2005; Hilu et al., 2003); however, the placement of these taxa has varied. *Xdh* data are mostly in agreement with results obtained from previous studies, although the *Xdh* trees are most closely similar to those of Soltis et al. (2005). Ranunculales

and Proteaceae (72%), Sabiaceae (35%), Trochodendraceae (71%), Buxaceae (38%), and Gunneraceae (72%) form a grade in the *Xdh* likelihood tree (Fig. 3). Ranunculales and Proteaceae were sister, forming a clade with low support (70%). This clade was not found by Soltis et al. (2005) and could be a result of the *Xdh* analysis having low sampling of Proteaceae, including only a single representative, *Platanus* L.

Eschscholzia Cham. (Papaveraceae, 96%) and *Euptelea* Siebold. & Zucc. (Eupteleaceae, 50%) form a grade to the remaining Ranunculales, and the Ranunculales are sister to the remaining eudicots, consistent with other analyses (Chase et al., 1993; Soltis et al., 1997, 2000, 2003, 2005; Hoot et al., 1999). Lardizabalaceae is a strongly supported clade (100%) sister to the Berberidaceae and Ranunculaceae clade (99%). Berberidaceae and Ranunculaceae are strongly supported as monophyletic groups (100% and 92%, respectively). The placement of Dilleniaceae and Santalaceae, forming a clade that is sister to the asterids and rosids, also occurs in the MPTs. The

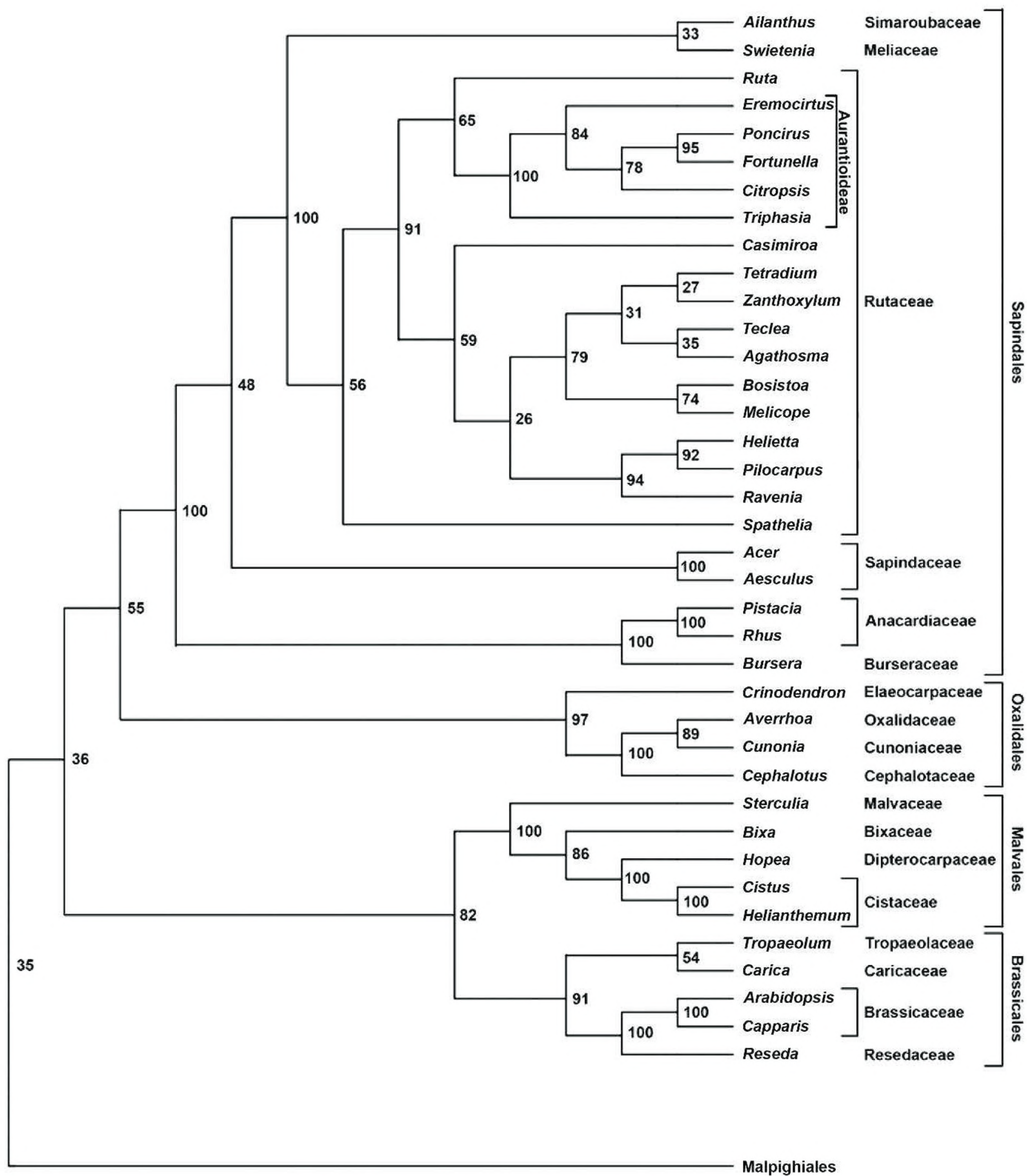


Figure 8. Bootstrap consensus tree of likelihood analyses among eurosid II. The numbers above the branches are likelihood bootstrap values.

sister-group relationship of *Dillenia* L. (Dilleniaceae) and *Osyris* L. (Santalaceae) (45%) has been found in other studies (Soltis et al., 2000, 2005), but the relationships of *Dillenia* are uncertain, with many different hypothesized relationships (compare Savolainen et al., 2000; Soltis et al., 2000, 2003, 2005; Hilu et al., 2003). The relationships of the Santalaceae are equally uncertain (compare Savolainen et al., 2000; Soltis et al., 2000, 2003; Hilu et al., 2003).

EUDICOTS CLADE I

Asterids

The asterid clade, sensu Olmstead et al. (1992, 1993) and Judd and Olmstead (2004), is monophyletic (58% bootstrap support) in the *Xdh* tree (Fig. 5)

and consists of the four major lineages Cornales, Ericales, and euasterid I and II. In the *Xdh* tree the Ericales–Cornales clade is strongly supported (97%, Fig. 4). The euasterids I and II (Fig. 5) formed poorly supported sister clades (63%). Although support values are difficult to compare, this topology is mostly in agreement with other multigene studies (Savolainen et al., 2000; Soltis et al., 2000).

Ericales–Cornales

The monophyly of Ericales received moderate support (83%) in the *Xdh* analysis, while the monophyly of the Cornales lacks bootstrap support (31%, Fig. 4). The resolution and support for the Ericales clade were similar to those in other analyses, including Bremer et al.’s (2002) six-gene plastid

analysis, Anderberg et al.'s (2002) five-gene analysis from plastid and mitochondrial regions, and Schönenberger et al.'s (2005) 11-gene analysis from nuclear, mitochondrial, and chloroplast regions. Within the Ericales, the *Xdh* analysis recovered four major clades, with multiple-sampled families monophyletic: (1) The first clade (100%) contains Tetrameristaceae and Balsaminaceae sister to Marcgraviaceae. (2) The second clade (100%) consists of the Primulaceae. (3) The third clade is tentative due to low support (18%) and has Sapotaceae sister to Lecythidaceae, which is sister to a Fouquieriaceae–Ebenaceae clade. The lack of samples from Polemoniaceae, which typically has been found to be sister to Fouquieriaceae (Savolainen et al., 2000 [*atpB*, *atpB/rbcL*]; Anderberg et al., 2002; Bremer et al., 2002; Soltis et al., 2005), is problematic in this analysis and may explain the low support for this group. (4) The fourth clade, also with low support, has two major lineages. The first contains Theaceae, sister to a clade containing Symplocaceae, and Styracaceae, sister to Diapensiaceae. The second is weakly supported and contains Sarraceniaceae sister to a clade consisting of Pentaphragmaceae, Theaceae, Actinidiaceae, Cyrillaceae, Clethraceae, and Ericaceae. Other multigene studies have found Sarraceniaceae sister to Actinidiaceae, together sister to Clethraceae, Cyrillaceae, and Ericaceae (Anderberg et al., 2002; Bremer et al., 2002). *Xdh* provides a similar inference although with varying support values. Using morphological data, Judd and Kron (1993) found the same topology for relationships with the core Ericales as found in the *Xdh* analysis.

Euasterid I (Lamids)

The euasterid I clade (consisting of Boraginaceae, Garryales, Gentianales, Lamiales, and Solanales) has strong bootstrap support (88%, Fig. 5) in the *Xdh* analysis, whereas most other studies have shown low to moderate support levels for it (Savolainen et al., 2000 [*atpB*, *rbcL*, *atpB/rbcL*]; Soltis et al., 2000, 2003; Albach et al., 2001; Bremer et al., 2002; Hilu et al., 2003). The first branching lineage is the Garryales and this is in agreement with previous studies (Savolainen et al., 2000 [*atpB*, *rbcL/atpB*]; Soltis et al., 2000, 2003; Albach et al., 2001). There is strong support (98%) for the relationships found among Boraginaceae, Gentianales, Solanales, and Lamiales. Monophyly of the orders Gentianales (100%) and Lamiales (including *Borago* L., 97%) also obtained strong support. This level of support has been found in most other studies (Savolainen et al., 2000 [*atpB*, *rbcL*, *atpB/rbcL*]; Soltis et al., 2000; Albach et al., 2001; Bremer et al., 2002; Hilu et al.,

2003). The Gentianales has Rubiaceae as sister to members of the Loganiaceae, Apocynaceae, and Gelsemiaceae. The Lamiales contains representative taxa from Bignoniaceae, Boraginaceae, Calceolariaceae, Gesneriaceae, and Plantaginaceae. In a clade representing members from the Plantaginaceae and Boraginaceae, *Erinus* L. (Plantaginaceae) is sister to *Borago* (Boraginaceae).

Euasterid II (Campanulids)

The euasterid II clade is weakly supported (50%) and consists of Aquifoliales, Asterales, Bruniaceae, Dipsacales, Escalloniaceae, and Apiales (Fig. 5). Aquifoliales as the first branching lineage is in agreement with most previous studies (Chase et al., 1993; Savolainen et al., 2000 [*atpB*, *atpB/rbcL*]; Soltis et al., 2000, 2003, 2005; Albach et al., 2001; Bremer et al., 2002) excluding those performed by *matK* analysis (Hilu et al., 2003), in which Aquifoliales aligned with the euasterid I (Campanulids) clade. Most analyses agree that euasterid II is composed of four major orders (Aquifoliales, Apicales, Dipsacales, and Asterales); however, the sister-group relationships differ among analyses. The most recent multigene analysis of Burleigh et al. (2009) is in agreement with this study in having Aquifoliales basal and sister to Asterales, together sister to Apiales and Dipsacales. The Apiales consists of approximately 10 families, the Asterales consists of 12 families, and the Aquifoliales contains approximately five families. The *Xdh* study sampled eight families of the possible 27 from these three orders; additional samples from the remaining families will most likely increase the internal resolution and provide additional support within euasterid II.

EUDICOTS CLADE II

The eudicots clade II likelihood tree contains the following taxa: Saxifragales, Myrtales–Caryophyllales–Cucurbitales, Crossosomatales, Geraniales, eurosid I (without Cucurbitales and Oxalidales), and eurosid II, with Oxalidales recovered in the *Xdh* tree. The MPT found Saxifragales, Crossosomatales, Geraniales, eurosid I (with Cucurbitales, Fabales, Fagales, Myrtales, and Rosales), and eurosid II (with Brassicales, Celastrales, Malpighiales, Malvales, Oxalidales, and Sapindales) as groups.

Saxifragales

The *Xdh* analysis has low support (73%) for the monophyly of Saxifragales (Fig. 6). Three well-supported clades are resolved within this order. The

first clade includes Iteaceae sister to Grossulariaceae (100%). A similar topology was recovered by Fishbein et al. (2001), Hilu et al. (2003), and Soltis et al. (2005). The second clade consists of a monophyletic Hamamelidaceae (100%), which has been documented in many other studies (Savolainen et al., 2000 [*atpB/rbcL*]; Soltis et al., 2000, 2003, 2005; Fishbein et al., 2001; Hilu et al., 2003). A third clade including Altingiaceae, Cercidiphyllaceae, Daphniphyllaceae, Hamamelidaceae, and Paeoniaceae is well supported (100%). Several major families are missing, such as Saxifragaceae and Crassulaceae. Therefore, it is still premature to comment on the position of the taxa in the third clade. Soltis et al. (2007a) also were unable to recover stable relationships among the woody Saxifragales. The Saxifragales circumscription contains a wide range of morphological diversity that departs from traditional classification. The nonmolecular synapomorphies have not yet been found and more research is needed for clarification.

Myrtales–Caryophyllales–Cucurbitales

The analyses recover a weakly supported (6%) clade consisting of Myrtales, Caryophyllales, and Cucurbitales (Fig. 6). This inference may be misleading or incorrect due to the weak support because no other analyses have grouped these orders together. In contrast, the MPT is like previous MP analyses, revealing Caryophyllales as sister to the eudicots, and Cucurbitales as sister to Fagales. In a recent analysis, the Caryophyllales clade has been variously placed at the base of the asterids (Hilu et al., 2003; Soltis et al., 2005), sister to the Saxifragales (Soltis et al., 2003), sister to the eudicots (Savolainen et al., 2000 [*rbcL, atpB/rbcL*]; Soltis et al., 2000), or sister to the asterids and rosids (Savolainen et al., 2000 [*atpB*]). Cucurbitales are typically found associated with a nitrogen-fixing clade, which includes Rosales, Fagales, Fabales, and Cucurbitales (Soltis et al., 1995, 1997, 2000, 2003, 2005; Savolainen et al., 2000 [*atpB, rbcL, atpB/rbcL*]; Hilu et al., 2003). The position of the Myrtales is usually nested within, or sister to, a clade consisting of Malvales, Brassicales, and Sapindales (Soltis et al., 1997, 2000, 2005; Savolainen et al., 2000 [*atpB, rbcL, atpB/rbcL*]; Hilu et al., 2003). The Myrtales clade in the *Xdh* data set has two members with partial sequences that may be creating errors in their placement. In all three orders, the families are underrepresented; we sampled only two of the approximately 13 families of the Myrtales, seven of the approximately 29 families of the Caryophyllales, and three of the approximately seven families of the

Cucurbitales. The differences between the likelihood and MP analyses need to be further explored using all genes in independent and combined analyses. Both factors, partial sequences and underrepresentation of families, are most likely contributing to this unusual placement of the Myrtales.

Two families were sampled for the Myrtales (63%), Myrtaceae and Penaeaceae. *Metrosideros* Banks ex Gaertn., *Myrciaria* O. Berg, and *Psidium* L. formed a monophyletic Myrtaceae with low support (63%, Fig. 6). This order has had various circumscriptions in the past, consisting of a number of families, and it would be premature to speculate on its composition, as well as relationships within the order and with other rosids. The monophyly of this order is in agreement with the results of Conti et al. (1996) based only on *rbcL*, Savolainen et al. (2000) based on *atpB/rbcL*, Soltis et al. (2000, 2005) based on *rbcL/atpB/18S*, and Hilu et al. (2003) based on *matK*.

The *Xdh* sequence data strongly support the circumscription of the Caryophyllales (95%, Fig. 6). The monophyly of this order has been indicated in previous molecular studies (Savolainen et al., 2000 [*atpB, rbcL, atpB/rbcL*]; Soltis et al., 2000, 2003, 2005; Cuénoud et al., 2002; Hilu et al., 2003). Two clades within Caryophyllales are recognized with *Xdh* data. Caryophyllales I (100%) includes Caryophyllaceae and Nyctaginaceae, followed by a well-supported Portulacaceae grade, with Cactaceae nested within (100%). Caryophyllales II (39%) contains a clade of the carnivorous families Droseraceae and Nepenthaceae (98%) sister to a clade of Plumbaginaceae and Simmondsiaceae (57%). The position of the Simmondsiaceae has varied, being included in the Caryophyllales II clade in some analyses (Savolainen et al., 2000 [*rbcL/atpB*]), while, in others, it has been included in the Caryophyllales I clade (Soltis et al., 2000, 2005; Cuénoud et al., 2002; Hilu et al., 2003). One of the most comprehensive studies of the Caryophyllales, by Cuénoud et al. (2002), indicates that the position of *Simmondsia* Nutt. lacks strong support, demonstrating that additional work is needed. Our results show Simmondsiaceae as part of the Caryophyllales II clade with strong support. Members of Caryophyllales have some non-DNA features supporting its basic circumscription, but many gaps still remain to be filled (for characters, see Nandi et al., 1998; Stevens, 2001, onwards).

The Cucurbitales were defined by the Angiosperm Phylogeny Group (1998) and Angiosperm Phylogeny Group II (2003) as containing the following seven families: Anisophylleaceae, Begoniaceae, Coriariaceae, Corynocarpaceae, Cucurbitaceae, Datisceae, and Tetramelaceae. Begoniaceae, Coriariaceae, and

Cucurbitaceae were included in this analysis (Fig. 6), which found the Cucurbitales to be monophyletic, although this was only weakly supported (72%). Other gene studies have also found the Cucurbitales to be monophyletic (Savolainen et al., 2000 [*atpB*, *rbcL*, *atpB/rbcL*]; Soltis et al., 2000, 2003, 2005; Hilu et al., 2003; Zhang et al., 2006). Morphological and phytochemical synapomorphies for the Cucurbitales are still unclear; even the most recent study by Zhang et al. (2006), using 13 molecular markers and morphological characters, could not identify a synapomorphy for the order.

Crossosomatales

The Crossosomatales clade, here consisting of Staphyleaceae and Stachyuraceae, is well supported (100%, Fig. 6). Both the *rbcL/atpB* analysis (Savolainen et al., 2000) and the analysis of 18S, *rbcL*, and *atpB* (Soltis et al., 2000) contained a well-supported Crossosomatales clade. Stachyuraceae has previously been placed in the Violales, whereas Staphyleaceae has been placed in the Sapindales (Cronquist, 1981). Thus, another group of families previously placed in distantly related orders come together in a well-supported clade. Matthews and Endress (2005) investigated the floral characters and revealed potential new synapomorphies for a close relationship among the core member of Crossosomatales. Among the prominent floral features of Crossosomatales are solitary flowers, presence of a floral cup, imbricate sepals with outermost smaller than inner, pollen grains with horizontally extended endoapertures, shortly stalked gynoecium, postgenitally united carpel tips forming a compitum, stigmatic papillae 2-cellular or more, ovary locules tapering upward, long integuments forming zigzag micropyles, cell clusters with bundles of long yellow crystals, mucilage cells, seeds with smooth, sclerified testa and without differentiated tegmen. However, more molecular and morphological work is needed to resolve both the composition of this clade and its position within the rosids.

Geraniales

Two taxa of Geraniales were sampled for the *Xdh* analysis, *Geranium* L. (Geraniaceae) and *Greyia* Hook. & Harv. (Melianthaceae) (Fig. 6). These two families are held together in a weakly supported clade (61%). These families were previously placed (Cronquist, 1981; Takhtajan, 1997) in different orders, and no known morphological synapomorphies hold the present order together. Geraniales have been found in previous analyses (Savolainen et al., 2000

[*atpB/rbcL*]; Soltis et al., 2000) to be sister to the Crossosomatales. While there is no support for that sister relationship here, likelihood values at the base of the grades leading up to the eurosids are low (12% or less) and rearrangements may be expected with additional sampling of taxa and inclusion/combination with more data sets.

Eurosid I

In the analysis, the nitrogen-fixing clade (including Rosales, Fagales, Fabales, and Cucurbitales; Soltis et al., 1995), excluding Cucurbitales, is weakly supported at the base of the eurosid I grade (20%), followed by Celastrales (88%) and Malpighiales (< 50%) (Fig. 7). Three nitrogen-fixing subclades received low to strong support and have no support as a clade (28%): Fabales (87%), Fagales (100%), and Rosales (65%). These subclades have been placed together with the order Cucurbitales in other single and multigene analyses (Savolainen et al., 2000 [*atpB*, *rbcL*, *atpB/rbcL*]; Soltis et al., 2000, 2003; Hilu et al., 2003). The Fabales consist of well-supported Fabaceae (100%) and Quillajaceae. The order is sister to the Fagales, which contains a mostly well-supported clade of Fagaceae, Betulaceae, Myricaceae, and Juglandaceae. The Rosales are sister to Fabales and Fagales. This order consists of Rosaceae sister to the rest of the order followed by Rhamnaceae sister to *Humulus* L. (Cannabaceae), Ulmaceae sister to *Celtis* L. (Cannabaceae), and Moraceae sister to Urticaceae. The Cannabaceae are not monophyletic in the *Xdh* analysis because they group with members of the Rhamnaceae and Ulmaceae; however, the placement of *Humulus* is based on a partial sequence, which may have caused anomalies. Soltis et al. (2005) found similar clades including a nonmonophyletic Cannabaceae in the Rosaceae. As mentioned earlier, the MP analysis found eurosid I to consist of Cucurbitales, Fabales, Fagales, Myrtales, and Rosales.

The *Xdh* analysis supports (100%) the Celastrales clade (100%, Fig. 7). Most recent analyses support the order Celastrales, including various families placed next to the Malpighiales (Zhang & Simmons, 2006), Oxalidales (Hilu et al., 2003; Soltis et al., 2005), or the Oxalidales and Malpighiales (Soltis et al., 2000). The *Xdh* likelihood and MP analysis showed Celastrales in a well-supported clade sister to the Malpighiales and eurosid II clades (88%).

The Malpighiales form a weakly supported clade with three supported subclades: Salicaceae (100%), Malpighiaceae (100%), and part of the Euphorbiaceae (73%, Fig. 7). The Euphorbiaceae are not monophyletic, as also found in the analyses by Soltis

et al. (2000) and Tokuoka and Tobe (2006, four-marker study). In the *Xdh* analysis, *Codiaeum variegatum* (L.) Rumph. ex A. Juss., *Euphorbia neriifolia* L., and *E. pseudocactus* A. Berger of the Euphorbiaceae formed a clade, whereas *Phyllanthus flexuosus* Müll. Arg. (Euphorbiaceae—now Phyllanthaceae) formed a sister grouping with *Passiflora suberosa* L. (Passifloraceae). This study sampled the following seven families: Chrysobalanaceae, Clusiaceae, Euphorbiaceae, Malpighiaceae, Passifloraceae, Salicaceae, and Violaceae. Previous analyses have found that the Malpighiales encompass at least 28 morphologically diverse families (Soltis et al., 2000). The relationships among these families are poorly understood and further studies incorporating additional taxa and non-DNA characters are needed to resolve relationships within this heterogeneous order.

Eurosid II (including Oxalidales)

The *Xdh* analysis produced a very weakly supported (< 50%) eurosid II clade, including members of the Brassicales, Malvales, Oxalidales, and Sapindales (Fig. 8). Each of these orders was recovered in the *Xdh* tree with strong support: Brassicales (91%), Malvales (100%), Oxalidales (97%), and Sapindales (100%). The *Xdh* data provided support for a sister-group relationship of Brassicales and Malvales (82%). This relationship was also obtained in *rbcL/atpB* analyses but was not well supported (Savolainen et al., 2000), although in the *matK* analysis these orders were strongly supported (89%, Hilu et al., 2003). In a family-level analysis examining glucosinolate biosynthesis taxa using *rbcL* and 18S, this sister-group relationship of Brassicales and Malvales was also found (Rodman et al., 1998). A clade consisting of the Sapindales, Brassicales, and Malvales has been found in some other recent analyses, but typically this association has not been well supported (Savolainen et al., 2000 [*rbcL/atpB*]; Soltis et al., 2000, 2003; Hilu et al., 2003). In Soltis et al.'s (2000, 2003) three- and four-gene analyses, weak to moderate support was found for Sapindales being sister to Malvales, whereas no support was found for an association with Brassicales.

This study included representatives of Brassicaceae, Caricaceae, Resedaceae, and Tropaeolaceae of Brassicales (Fig. 8). *Arabidopsis* and *Capparis* L. (Brassicaceae) formed a well-supported clade (100%) sister to *Reseda* L. (Resedaceae, 100%), whereas *Tropaeolum* L. (Tropaeolaceae) and *Carica* L. (Caricaceae) formed a weakly supported clade (54%). Chemical compounds including glucosinolates from phenylalanine and/or tyrosine and myricetin and other methylated flavonols may provide

synapomorphies for the order (Stevens, 2001, onward).

The Malvales comprised a well-supported (100%) order consisting of *Sterculia* L. (Malvaceae), *Bixa* L. (Bixaceae) (86%), *Hopea* Roxb. (Dipterocarpaceae) (100%), and *Cistus* L. and *Helianthemum* Mill. (Cistaceae) (100%, Fig. 8). It is premature to speculate on the familial relationships of Brassicales and Malvales using *Xdh* until additional taxa are sequenced.

The Sapindales are well supported (100%), with a number of well-supported subclades: (1) a strongly supported (100%) sister group of Anacardiaceae (100%) and Burseraceae; (2) a well-supported (100%) Sapindaceae with *Aesculus* L. and *Acer* L.; (3) a well-supported clade (100%) of Meliaceae, Simaroubaceae, and Rutaceae; and (4) a well-supported Rutaceae (91%), excluding *Spathelia* L. with the Aurantioideae (100%) and *Ruta* L. basal (65%) to the Aurantioideae, as has been previously found (Chase et al., 1999). Secondary compounds such as the ethereal oils and myricetin provide synapomorphies for the order (Morton, unpublished data).

Oxalidales has been positioned as sister to Celastrales and Huaceae (< 50%) (Savolainen et al., 2000 [*atpB/rbcL*]), Celastrales (Hilu et al., 2003; Soltis et al., 2005), and Malpighiales (Soltis et al., 2000 [*< 50%*], 2003 [*< 50%*]). The inclusion of the Oxalidales as sister to the Sapindales (56%) in the *Xdh* analysis is unique but not well supported (56%). The MP analysis, however, supports Celastrales and Malpighiales sister to Oxalidales, and this clade is sister to a Brassicales–Malvales–Sapindales clade. The phylogenetic analyses of Soltis et al. (2005) indicated that Oxalidales consists of six families: Brunelliaceae, Cephalotaceae, Connaraceae, Cunoniaceae, Elaeocarpaceae, and Oxalidaceae. This study included a well-supported clade of Cephalotaceae, Cunoniaceae, Elaeocarpaceae, and Oxalidaceae (97%). These families have all been previously placed in different orders, Cephalotaceae and Cunoniaceae in Rosales, Elaeocarpaceae in Malvales, and Oxalidaceae in Geraniales (Cronquist, 1981). Thus, here the Oxalidales comprise taxa that were once considered distantly related. Currently, no morphological synapomorphies are evident for eurosid II with or without Oxalidales (Judd et al., 1994, 2007).

The rosids remain inadequately understood as to supported phylogenetic positions, non-DNA synapomorphies for major clades, and patterns of character evolution. Most of these groups contain unexpected assemblages of taxa that previously have been placed

in distantly related orders. For these new groupings, many gaps in the non-DNA information render many features phylogenetically uninformative.

CONCLUSION

The *Xdh* tree of 247 genera and 1265 bp produced a likelihood tree that is robust and mostly congruent with other single or multigene studies (Savolainen et al., 2000; Soltis et al., 2000, 2003, 2005; Hilu et al., 2003; Qiu et al., 2006). The potential utility of *Xdh* in angiosperm phylogenetics can be gauged by the number of genera and base pairs used to produce this topology in comparison to other single and multigene analyses in which substantially more genera and base pairs were used to produce similar topologies and support levels. The *Xdh* topology demonstrated 71% of the nodes with bootstrap support greater than 50% in comparison to the *matK* topology (374 taxa), with 79%, and the reported 7%–24% for individual analyses of *rbcL*, *atpB* (357 taxa), and 18S (223 taxa) ribosomal DNA clades (Hilu et al., 2003: table 2, 1761). The number of nodes receiving greater than 50% support in the *Xdh* analysis was in the same range as was obtained for the *matK* and the combined analyses of three to four genes from two genomic regions. The number of parsimony-informative sites (1068–1069) was similar to the *matK* analysis (1083–1062).

Xdh analysis provides good support values for monophyly of the Caryophyllales, Ericales, and Cornales, euasterids I (lamids), Magnoliales and Laurales, Malvales, Rutaceae, Oxalidales, Brassicales, and Sapindales.

As stated earlier, paralogy exists for all genes. More than one copy was found in selected taxa; however, most had high sequence similarity, suggesting recent duplication. Among two taxa that had lower values, the highly divergent sequence was easily recognized. When the divergent sequence was removed, the remaining sequences had 98%–99% homology. Having more than one gene phylogeny to compare the angiosperm phylogeny is a good indicator of paralogy. The likelihood and MP topology of the *Xdh* gene are comparable to other individual gene data sets. The potential utility of this gene is quite promising at the ordinal, familial, and generic levels.

Although phylogenetic relationships of angiosperms have been explored using a number of gene sequences from the chloroplast and mitochondria, the nucleus remains poorly explored. Phylogenetics has achieved some understanding of angiosperm relationships, and the present study contributes to our knowledge. There are still parts of the tree that

remain unresolved or unsupported by both molecular and morphological data. As Chase et al. (2006) discovered, adding mitochondrial and nuclear genes to four plastid genes produced a highly congruent result that is an improvement over their predecessors. We must continue to combine chloroplast, mitochondrial, nuclear, and morphological data to improve our understanding of angiosperm evolution.

Literature Cited

- Albach, D. C., P. S. Soltis, D. E. Soltis & R. G. Olmstead. 2001. Phylogenetic analysis of asterids based on sequences of four genes. *Ann. Missouri Bot. Gard.* 88: 163–212.
- Anderberg, A. A., C. Rydin & M. Källersjö. 2002. Phylogenetic relationships in the order Ericales s.l.: Analyses of molecular data from five genes from the plastid and mitochondrial genomes. *Amer. J. Bot.* 89: 677–687.
- Aoki, S., K. Uehara, M. Imafuku, M. Hasebe & M. Ito. 2004. Phylogeny and divergence of basal angiosperms inferred from *APETALA3*- and *PISTILLATA*-like MADS-box genes. *J. Pl. Res.* 117: 229–244.
- Angiosperm Phylogeny Group. 1998. An ordinal classification for the families of flowering plants. *Ann. Missouri Bot. Gard.* 85: 531–553.
- Angiosperm Phylogeny Group II. 2003. An update of Angiosperm Phylogeny Group classification for the orders and families of flowering plants. *Bot. J. Linn. Soc.* 141: 399–436.
- Barkman, T. J., G. Chenery, J. R. McNeal, J. Lyons-Weiler, W. J. Ellisens, G. Moore, A. D. Wolfe & C. W. dePamphilis. 2000. Independent and combined analyses of sequences from all three genomic compartments converge on the root of flowering plant phylogeny. *Proc. Natl. Acad. Sci. U.S.A.* 97: 13166–13171.
- Borsch, T., K. W. Hilu, D. Quandt, V. Wilde, C. Neinhuis & W. Barthlott. 2003. Noncoding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *J. Evol. Biol.* 16: 558–576.
- Borsch, T., C. Löhne, K. Müller, K. W. Hilu, S. Wanke, A. Worberg, W. Barthlott, C. Neinhuis & D. Quandt. 2005. Toward understanding basal angiosperm diversification: Recent insights using rapidly evolving genome regions. *Nova Acta Leopold.* 92: 85–110.
- Bremer, B., K. Bremer, N. Heidari, P. Erixon, R. G. Olmstead, A. A. Anderberg, M. Källersjö & E. Barkhordarian. 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. *Molec. Phylogen. Evol.* 24: 274–301.
- Burleigh, J. G., K. W. Hilu & D. E. Soltis. 2009. Inferring phylogenies with incomplete data sets: A 5-gene, 567-taxon analysis of angiosperms. *BMC Evol. Biol.* 9: 61.
- Cantino, P. D., J. A. Doyle, S. W. Graham, W. S. Judd, R. G. Olmstead, D. E. Soltis, P. S. Soltis & M. J. Donoghue. 2007. Towards a phylogenetic nomenclature of *Tracheophyta*. *Taxon* 56: 822–846.
- Chanderbali, A. S., H. van der Werff & S. S. Renner. 2001. Phylogeny and historical biogeography of Lauraceae: Evidence from the chloroplast and nuclear genomes. *Ann. Missouri Bot. Gard.* 88: 104–134.
- Chase, M. W., D. E. Soltis, R. G. Olmstead, D. Morgan, D. H. Les, B. D. Mishler, M. R. Duvall, R. A. Price, H. G. Hills, Y.-L. Qiu, K. A. Kron, J. H. Rettig, E. Conti, J. D.

- Palmer, J. R. Manhart, K. J. Sytsma, H. J. Michaels, W. J. Kress, K. G. Karol, W. D. Clark, M. Hedrén, B. S. Gaut, R. K. Jansen, K.-J. Kim, C. F. Wimpee, J. F. Smith, G. R. Furnier, S. H. Strauss, Q.-Y. Xiang, G. M. Plunkett, P. S. Soltis, S. M. Swensen, S. E. Williams, P. A. Gadek, C. J. Quinn, L. E. Eguiarte, E. Golenberg, G. H. Learn Jr., S. W. Graham, S. C. H. Barrett, S. Dayanandan & V. A. Albert. 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann. Missouri Bot. Gard.* 80: 528–580.
- Chase, M. W., C. M. Morton & J. A. Kallunki. 1999. Phylogenetic relationships of Rutaceae: A cladistic analysis of the subfamilies using evidence from RBC and ATP sequence variation. *Amer. J. Bot.* 86: 1191–1199.
- Chase, M. W., D. E. Soltis, P. S. Soltis, P. J. Rudall, M. F. Fay, W. J. Hahn, S. Sullivan, J. Joseph, M. Molvray, P. J. Kores, T. J. Givnish, K. J. Sytsma & J. C. Pires. 2000. Higher-level systematics of the monocotyledons: An assessment of current knowledge and a new classification. Pp. 3–16 in K. K. L. Wilson & D. Morrison (editors), *Proceedings of Monocots II: The Second International Symposium on the Comparative Biology of the Monocotyledons*. CSIRO Press, Sydney.
- Chase, M. W., M. F. Fay, D. S. Devey, O. Maurin, N. Rønsted, T. J. Davies, Y. Pillon, G. Petersen, O. Seberg, M. N. Tamura, C. B. Asmussen, K. Hilu, T. Borsch, J. I. Davis, D. W. Stevenson, J. C. Pires, T. J. Givnish, K. J. Sytsma, M. A. McPherson, S. W. Graham & H. S. Rai. 2006. Multigene analyses of monocot relationships: A summary. *Aliso* 22: 63–75.
- Conti, E., A. Litt & K. J. Sytsma. 1996. Circumscription of Myrtales and their relationships to other rosids: Evidence from *rbcL* sequence data. *Amer. J. Bot.* 83: 221–223.
- Cronquist, A. 1981. *An Integrated System of Classification of Flowering Plants*. Columbia University Press, New York.
- Cuénoud, P., V. Savolainen, L. W. Chatrou, M. Powell, R. J. Grayer & M. W. Chase. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. *Amer. J. Bot.* 89: 132–144.
- Doyle, J. A. & P. K. Endress. 2000. Morphological phylogenetic analysis of basal angiosperms: Comparison and combination with molecular data. *Int. J. Pl. Sci.* 161: S121–S153.
- Doyle, J. A., H. Eklund & P. S. Herendeen. 2003. Floral evolution in Chloranthaceae: Implications of a morphological phylogenetic analysis. *Int. J. Pl. Sci.* 164: S365–S382.
- Fishbein, M., C. Hibsich-Jetter, D. E. Soltis & L. Hufford. 2001. Phylogeny of Saxifragales (angiosperms, eudicots): Analysis of a rapid, ancient radiation. *Syst. Biol.* 50: 817–847.
- Friis, E. M., P. R. Crane & K. R. Pedersen. 1986. Floral evidence for Cretaceous chloranthoid angiosperms. *Nature* 320: 163–164.
- Goremykin, V. V., K. I. Hirsch-Ernst, S. Wolfl & F. H. Hellwig. 2003. Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. *Molec. Biol. Evol.* 20: 1499–1505.
- Graham, S. W. & R. G. Olmstead. 2000. Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. *Amer. J. Bot.* 87: 1712–1730.
- Graham, S. W., J. M. Zgurski, M. A. McPherson, D. M. Cherniawsky, J. M. Saarela, E. F. C. Horne, S. Y. Smith, W. A. Wong, H. E. O'Brien, V. L. Biron, J. C. Pires, R. G. Olmstead, M. W. Chase & H. S. Rai. 2006. Robust inference of monocot deep phylogeny using an expanded multigene plastid data set. *Aliso* 22: 3–21.
- Hesberg, C., R. Hänsch, R. R. Mendel & F. Bittner. 2004. Tandem orientation of duplicated xanthine dehydrogenase genes from *Arabidopsis thaliana*: Differential gene expression and enzyme activities. *J. Biol. Chem.* 279: 13547–13554.
- Hilu, K. W., T. Borsch, K. Müller, D. E. Soltis, P. S. Soltis, V. Savolainen, M. W. Chase, M. P. Powell, L. A. Alice, R. Evans, H. Sauquet, C. Neinhuis, T. A. B. Slotta, J. G. Rohwer, C. S. Campbell & L. W. Chatrou. 2003. Angiosperm phylogeny based on *matK* sequence information. *Amer. J. Bot.* 90: 1758–1776.
- Holmquist, R. 1983. Transitions and transversions in evolutionary descent: An approach to understanding. *J. Molec. Evol.* 19: 134–144.
- Hoot, S. B., S. Magallon & P. R. Crane. 1999. Phylogeny of basal eudicots based on three molecular data sets: *atpB*, *rbcL*, and 18S nuclear ribosomal DNA sequences. *Ann. Missouri Bot. Gard.* 86: 1–32.
- Jansen, R. K., Z. Cai, L. A. Raubeson, H. Daniell, C. W. dePamphilis, J. Leebens-Mack, K. F. Müller, M. Guisinger-Bellian, R. C. Haberle, A. K. Hansen, T. W. Chumley, S.-B. Lee, R. Perry, J. R. McNeal, J. V. Kuehl & J. L. Boore. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc. Natl. Acad. Sci. U.S.A.* 104: 19369–19374.
- Judd, W. S. & K. A. Kron. 1993. Circumscription of Ericaceae (Ericales) as determined by preliminary cladistic analyses based on morphological, anatomical, and embryological features. *Brittonia* 45: 99–114.
- Judd, W. S. & R. G. Olmstead. 2004. A survey of tricolpate (eudicot) phylogenetic relationships. *Amer. J. Bot.* 91: 1627–1644.
- Judd, W. S., R. W. Sanders & M. J. Donoghue. 1994. Angiosperm family pairs: Preliminary phylogenetic analyses. *Harvard Pap. Bot.* 5: 1–51.
- Judd, W. S., C. S. Campbell, E. A. Kellogg, P. F. Stevens & M. J. Donoghue. 2007. *Plant Systematics: A Phylogenetic Approach*, 3rd ed. Sinauer Associates, Sunderland, Massachusetts.
- Löhne, C. & T. Borsch. 2005. Molecular evolution and phylogenetic utility of the *petD* group II intron: A case study in basal angiosperms. *Molec. Biol. Evol.* 22: 317–332.
- Mathews, S. & M. J. Donoghue. 1999. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* 286: 947–950.
- Mathews, S. & M. J. Donoghue. 2000. Basal angiosperm phylogeny inferred from duplicated phytochromes A and C. *Int. J. Pl. Sci.* 161: S41–S55.
- Matthews, M. L. & P. K. Endress. 2005. Comparative floral structure and systematics in Crossosomatales (Crossosomataceae, Stachyuraceae, Staphyleaceae, Aphloiaceae, Geissolomataceae, Ixerbaceae, Strasburgeriaceae). *Bot. J. Linn. Soc.* 147: 1–46.
- Mendel, R. R. & R. Hänsch. 2002. Molybdoenzymes and molybdenum cofactor in plants. *J. Exp. Bot.* 53: 1689–1698.
- Moore, M. J., C. D. Bell, P. S. Soltis & D. E. Soltis. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc. Natl. Acad. Sci. U.S.A.* 104: 19363–19368.
- Nandi, O. I., M. W. Chase & P. K. Endress. 1998. A combined cladistic analysis of angiosperms using *rbcL*

- and non-molecular data sets. *Ann. Missouri Bot. Gard.* 85: 137–214.
- Olmstead, R. G., H. J. Michaels, K. M. Scott & J. D. Palmer. 1992. Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbcL*. *Ann. Missouri Bot. Gard.* 79: 249–265.
- Olmstead, R. G., B. Bremer, K. M. Scott & J. D. Palmer. 1993. A parsimony analysis of the Asteridae s.l. based on *rbcL* sequences. *Ann. Missouri Bot. Gard.* 80: 700–722.
- Parkinson, C. L., K. L. Adams & J. D. Palmer. 1999. Multigene analyses identify the three earliest lineages of extant flowering plants. *Curr. Biol.* 9: 1485–1488.
- Pedersen, K. R., P. R. Crane, A. N. Drinnan & E. M. Friis. 1991. Fruits from the mid-Cretaceous of North America with pollen grains of the *Clavatipollenites* types. *Grana* 30: 577–590.
- Peery, R., K. R. Wilcox, C. Morton & L. A. Raubeson. 2008. Exploring the utility of the nuclear XDH gene for gymnosperm phylogenetics. *Botanical Society of America Abstract* 667.
- Posada, D. & K. A. Crandall. 1998. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
- Qiu, Y.-L., J. Lee, F. Bernasconi-Quadroni, D. E. Soltis, P. S. Soltis, M. Zanis, E. A. Zimmer, Z. Chen, V. Savolainen & M. W. Chase. 1999. The earliest angiosperms: Evidence from mitochondrial, plastid, and nuclear genomes. *Nature* 402: 404–407.
- Qiu, Y.-L., J. Lee, F. Bernasconi-Quadroni, D. E. Soltis, P. S. Soltis, M. Zanis, E. A. Zimmer, Z. Chen, V. Savolainen & M. W. Chase. 2000. Phylogeny of basal angiosperms: Analyses of five genes from three genomes. *Int. J. Pl. Sci.* 161: S3–S27.
- Qiu, Y.-L., O. Dombrovskaya, J. Lee, L. Li, B. A. Whitlock, F. Bernasconi-Quadroni, J. S. Rest, C. C. Davis, T. Borsch, K. W. Hilu, S. S. Renner, D. E. Soltis, P. S. Soltis, M. J. Zanis, J. J. Cannone, R. R. Gutell, M. Powell, V. Savolainen, L. W. Chatrou & M. W. Chase. 2005. Phylogenetic analysis of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. *Int. J. Pl. Sci.* 166: 815–842.
- Qiu, Y.-L., L. Libo, T. A. Hendry, R. Li, D. W. Taylor, M. J. Issa, A. J. Ronen, M. L. Vekaria & A. M. White. 2006. Reconstructing the basal angiosperm phylogeny: Evaluating information content of mitochondrial genes. *Taxon* 55: 837–856.
- Rodman, J. E., P. S. Soltis, D. E. Soltis, K. J. Sytsma & K. G. Karol. 1998. Parallel evolution of glucosinolate biosynthesis inferred from congruent nuclear and plastid gene phylogenies. *Amer. J. Bot.* 85: 997–1006.
- Rodríguez-Trelles, F., R. Tarrío & F. J. Ayala. 2003. Convergent neofunctionalization by positive Darwinian selection after ancient recurrent duplications of the *xanthine dehydrogenase* gene. *Proc. Natl. Acad. Sci. U.S.A.* 100: 13413–13417.
- Savolainen, V., M. W. Chase, S. B. Hoot, C. M. Morton, D. E. Soltis, C. Bayer, M. F. Fay, A. Y. de Bruijn, S. Sullivan & Y.-L. Qiu. 2000. Phylogenetics of flowering plants based upon a combined analysis of plastid *atpB* and *rbcL* gene sequences. *Syst. Biol.* 49: 306–362.
- Schönenberger, J., A. A. Anderberg & K. J. Sytsma. 2005. Molecular phylogenetics and patterns of floral evolution in the Ericales. *Int. J. Pl. Sci.* 166: 265–288.
- Soltis, D. E., P. S. Soltis, D. R. Morgan, S. M. Swensen, B. C. Mullin, J. M. Dowd & P. G. Martin. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc. Natl. Acad. Sci. U.S.A.* 92: 2647–2651.
- Soltis, D. E., P. S. Soltis, D. L. Nickrent, L. A. Johnson, W. J. Hahn, S. B. Hoot, J. A. Sweere, R. K. Kuzoff, K. A. Kron, M. W. Chase, S. M. Swensen, E. A. Zimmer, S.-M. Chaw, L. J. Gillespie, W. J. Kress & K. J. Sytsma. 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Ann. Missouri Bot. Gard.* 84: 1–49.
- Soltis, D. E., P. S. Soltis, M. W. Chase, M. E. Mort, D. C. Albach, M. Zanis, V. Savolainen, W. H. Hahn, S. B. Hoot, M. F. Fay, M. Axtell, S. M. Swensen, L. M. Prince, W. J. Kress, K. C. Nixon & J. S. Farris. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Bot. J. Linn. Soc.* 133: 381–461.
- Soltis, D. E., A. E. Senter, M. J. Zanis, S. Kim, J. D. Thompson, P. S. Soltis, L. P. Ronse De Craene, P. K. Endress & J. S. Farris. 2003. Gunnerales are sister to other core eudicots: Implications for the evolution of pentamery. *Amer. J. Bot.* 90: 461–470.
- Soltis, D. E., P. S. Soltis, P. K. Endress & M. W. Chase. 2005. *Phylogeny and Evolution of Angiosperms*. Sinauer Associates, Sunderland, Massachusetts.
- Soltis, D. E., J. W. Clayton, C. C. Davis, M. A. Gitzendanner, M. Cheek, V. Savolainen, M. Amorim & P. S. Soltis. 2007a. Monophyly and relationships of the enigmatic family Peridiscaceae. *Taxon* 56: 65–73.
- Soltis, D. E., M. A. Gitzendanner & P. S. Soltis. 2007b. A 567-taxon data set for angiosperms: The challenges posed by Bayesian analyses of large data sets. *Int. J. Pl. Sci.* 168: 137–157.
- Soltis, P. S., D. E. Soltis & M. W. Chase. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402: 402–404.
- Stefanović, S., D. W. Rice & J. D. Palmer. 2004. Long branch attraction, taxon sampling, and the earliest angiosperms: *Amborella* or monocots? *BMC Evol. Biol.* 4: 35.
- Stevens, P. F. 2001 (and onwards). Angiosperm Phylogeny Website. Version 8, June 2007. <<http://www.mobot.org/MOBOT/research/APweb/>>, accessed 2 December 2010.
- Swofford, D. L. 2004. *PAUP: Phylogenetic Analysis Using Parsimony*, Vers. 4.0b6. Sinauer Associates, Sunderland, Massachusetts.
- Takhtajan, A. 1997. *Diversity and Classification of Flowering Plants*. Columbia University Press, New York.
- Tokuoka, T. & H. Tobe. 2006. Phylogenetic analyses of Malpighiales using plastid, nuclear DNA sequences, with particular reference to the embryology of Euphorbiaceae sens. str. *J. Pl. Res.* 119: 599–616.
- Zanis, M. J., D. E. Soltis, P. S. Soltis, S. Mathews & M. J. Donoghue. 2002. The root of the angiosperms revisited. *Proc. Natl. Acad. Sci. U.S.A.* 99: 6848–6853.
- Zanis, M. J., P. S. Soltis, Y.-L. Qiu, E. Zimmer & D. E. Soltis. 2003. Phylogenetic analyses and perianth evolution in basal angiosperms. *Ann. Missouri Bot. Gard.* 90: 129–150.
- Zhang, L. B. & M. P. Simmons. 2006. Phylogeny and delimitation of the Celastrales inferred from nuclear and plastid genes. *Syst. Bot.* 31: 122–137.
- Zhang, L. B., M. P. Simmons, A. Kocyan & S. S. Renner. 2006. Phylogeny of the Cucurbitales based on DNA sequences of nine loci from three genomes: Implications for morphological and sexual system evolution. *Molec. Phylogen. Evol.* 39: 305–322.
- Zwickl, D. J. 2006. *Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets Under the Maximum Likelihood Criterion*. Ph.D. Dissertation, The University of Texas at Austin, Austin.

Appendix 1. Species sequenced for the present study, accession data, and GenBank accession numbers. See Materials and Methods. Order follows Stevens (2001 and onwards) and Group follows Burleigh et al. (2009).

Group	Order	Family	Species	Voucher/Source	Museum voucher/Number	GenBank number
Gymnosperms	Pinales	Podocarpaceae	<i>Podocarpus macrophyllus</i> (Thunb.) Sweet	X.16-33*A (FTBG)	<i>C. & A. Morton 2009-13</i> (CM)	EU264209
		Pinaceae	<i>Picea abies</i> (L.) H. Karst.	01/0181.A (PCPT)	<i>C. Morton 2010-1</i> (CM)	EU264211
			<i>Pinus elliotii</i> Engelm.	2002-0236*D (FTBG)	<i>C. Morton 2010-37</i> (CM)	EU264208
	Cycadales	Cycadaceae	<i>Cycas revoluta</i> Thunb.	99/0245 (PCPT)	<i>C. Morton 2010-2</i> (CM)	EU264212
	Ginkgoales	Ginkgoaceae	<i>Ginkgo biloba</i> L.	Morton	<i>C. Morton 2010-27</i> (CM)	EU264210
Angiosperms	Gnetales	Welwitschiaceae	<i>Welwitschia mirabilis</i> Hook. f.	University of Pittsburgh Biological Greenhouse	<i>C. Morton 2010-28</i> (CM)	EU264213
	Amborellales	Amborellaceae	<i>Amborella trichopoda</i> Baill.	2000.0348 (UC, UCBG)	2000.0348 (UC, UCBG)	EU264292
		Chloranthales	<i>Ascarina</i> J. R. Forst. & G. Forst., sp. indet.	<i>Thien 500</i> (NO)	<i>L. B. Thien 500</i> (NO)	EU264260
			<i>Chloranthus multistachys</i> C. Pei	<i>K. Wurdack 92-0010</i>	<i>K. Wurdack 92-0010</i> (US)	EU264259
	Nymphaeales	Nymphaeaceae	<i>Brasenia schreberi</i> J. F. Gmel.	<i>Qiu 91031</i> (NCU)	<i>Y.-L. Qiu 91031</i> (NCU)	EU264275
			<i>Nuphar lutea</i> (L.) Sm.	81.0800 (UCBG)	81.0800 (UCBG, UC)	EU264293
	Austrobaileyales	Austrobaileyaceae	<i>Nymphaea</i> L., sp. indet.	<i>Qiu 91029</i> (NCU)	<i>Y.-L. Qiu 91029</i> (NCU)	EU264262
			<i>Austrobaileya scandens</i> C. T. White	<i>Qiu 90030</i> (NCU)	<i>Y.-L. Qiu 90030</i> (NCU)	EU264258
			<i>Kadsura japonica</i> (L.) Dunal	72.0156 (UCBG)	72.0156 (UC, UCBG)	EU264261
			<i>Schisandra chinensis</i> (Turcz.) Baill.	70.0209 (UCBG)	70.0209 (UC, UCBG)	EU264271
Magnoliids	Ceratophyllales	Ceratophyllaceae	<i>Ceratophyllum demersum</i> L.	<i>C. R. Parks</i> (IND)	<i>C. R. Parks</i> (IND)	EU264214
		Canellaceae	<i>Canella winterana</i> (L.) Gaertn.	<i>Qiu 90017</i> (NCU)	<i>Y.-L. Qiu 90017</i> (NCU)	EU264255
	Canellales	Winteraceae	<i>Drimys winteri</i> J. R. Forst. & G. Forst.	<i>Qiu 90016</i> (NCU)	<i>Y.-L. Qiu 90016</i> (NCU)	EU264254
			<i>Tasmannia insipida</i> R. Br. ex DC.	<i>Qiu 90032</i> (NCU)	<i>Y.-L. Qiu 90032</i> (NCU)	EU264256
	Laurales	Atherospermataceae	<i>Atherosperma moschatum</i> Labill.	<i>Qiu 92007</i> (NCU)	<i>Y.-L. Qiu 92007</i> (NCU)	EU264257
			<i>Daphnandra micrantha</i> (Tul.) Benth.	<i>Whalen 132</i> (NCU)	<i>Whalen 132</i> (NCU)	EU264269
			<i>Doryphora sassafras</i> Endl.	<i>Qiu 98109</i> (Z)	<i>Y.-L. Qiu 98109</i> (Z)	EU264270
			<i>Calycanthus floridus</i> L.	65.0164 (UCBG)	65.0164 (UC, UCBG)	EU264294
	Gomortegaceae	Hernandiaceae	<i>Chimonanthus praecox</i> Lindl.	<i>Qiu 9</i> (NCU)	<i>Y.-L. Qiu 9</i> (NCU)	EU264268
			<i>Gomortega keule</i> (Molina) Baill.	2001.0213 (UCBG)	2001.0213 (UC, UCBG)	EU264287
	Lauraceae	Lauraceae	<i>Gyrocarpus americanus</i> Jacq.	96851*A (FTBG)	<i>C. & A. Morton 2009-17</i> (CM)	EU264302
			<i>Hernandia sonora</i> L.	DF311*B (FTBG)	<i>C. & A. Morton 2009-6</i> (CM)	EU264300
			<i>Cinnamomum zeylanicum</i> Blume	95/0619 (PCPT)	<i>C. Morton 2010-3</i> (CM)	EU264303
			<i>Cryptocarya meisneriana</i> Frodin	<i>Qiu 98048</i> (Z)	<i>Y.-L. Qiu 98048</i> (Z)	EU264272
	Monimiaceae	Monimiaceae	<i>Laurus nobilis</i> L.	<i>Qiu 94209</i> (IND)	<i>Y.-L. Qiu 94209</i> (IND)	EU264273
			<i>Hedycarya arborea</i> J. R. Forst. & G. Forst.	<i>Qiu 90028</i> (NCU)	<i>Y.-L. Qiu 90028</i> (NCU)	EU264276
			<i>Hortonia floribunda</i> Wight ex Arn.	<i>Qiu 02002</i> (no voucher)	<i>Y.-L. Qiu 02002</i> (no voucher)	EU264215
	Siparunaceae	Siparunaceae	<i>Kibara coriacea</i> (Blume) Hook. f. & Thomson	<i>M. Coode 7879</i> (K)	<i>M. Coode 7879</i> (K)	EU264301
			<i>Siparuna brasiliensis</i> A. DC.	<i>J. Lombardi 2714</i> (MO)	<i>J. Lombardi 2714</i> (MO)	EU264274

Appendix 1. Continued.

Group	Order	Family	Species	Voucher/Source	Museum voucher/Number	GenBank number
Monocots	Magnoliales	Annonaceae	<i>Annona glabra</i> L.	63271*B (FTBG)	<i>C. Morton 2010-38</i> (CM)	EU264298
			<i>Cananga odorata</i> (Lam.) Hook. f. & Thomson	2001-0325*A (FTBG)	<i>C. & A. Morton 2009-15</i> (CM)	EU264267
		Degeneriaceae	<i>Degeneria roseiflora</i> John M. Mill.	<i>J. M. Miller 1189-63</i> (K)	<i>J. M. Miller 1189-63</i> (K)	EU264266
		Eupomatiaceae	<i>Eupomatia bennettii</i> F. Muell.	86.0660 (UCBG)	86.0660 (UC, UCBG)	EU264299
		Himantandraceae	<i>Galbulimima belgegraveana</i> (F. Muell.) Sprague	<i>P. H. Weston 929</i> (NSW)	<i>P. H. Weston 929</i> (NSW)	EU264297
		Magnoliaceae	<i>Liriodendron tulipifera</i> L.	<i>Isaac 18669</i> (CM)	<i>J. & B. Isaac 18669</i> (CM)	EU264295
		Myristicaceae	<i>Magnolia acuminata</i> L.	<i>Isaac 18685</i> (CM)	<i>J. & B. Isaac 18685</i> (CM)	EU264296
			<i>Mauloutchia chapelieri</i> Warb.	<i>G. E. Schatz et al. 3847A</i> (MO)	<i>G. E. Schatz et al. 3847A</i> (MO)	EU264264
			<i>Myristica fragrans</i> Houtt.	<i>Qiu 92014</i> (NCU)	<i>Y.-L. Qiu 92014</i> (NCU)	EU264263
		Aristolochiaceae	<i>Aristolochia maxima</i> Jacq.	75602*A (FTBG)	<i>C. & A. Morton 2009-2</i> (CM)	EU264249
Piperales		Saururaceae	<i>Asarum canadense</i> L.	<i>Peter Kuhlman</i> (IND)	<i>P. Kuhlman</i> (IND)	EU264250
			<i>Thottea tomentosa</i> (Blume) Ding Hou	<i>M. W. Chase 1211</i> (K)	<i>M. W. Chase 1211</i> (K)	EU264251
			<i>Anemopsis californica</i> (Nutt.) Hook. & Arn.	<i>Qiu 97116</i> (IND)	<i>Y.-L. Qiu 97116</i> (IND)	EU264439
			<i>Houttuynia cordata</i> Thunb.	<i>Qiu 92016</i> (NCU)	<i>Y.-L. Qiu 92016</i> (NCU)	EU264253
			<i>Saururus cernuus</i> L.	78.0467 (UCBG)	78.0467 (UC, UCBG)	EU264252
			<i>Acorus gramineus</i> Sol. ex Aiton	94.1197 (UCBG)	94.1197 (UC, UCBG)	EU264285
		Araceae	<i>Orontium aquaticum</i> L.	<i>Qiu 97112</i> (IND)	<i>Y.-L. Qiu 97112</i> (IND)	EU264440
		Tofieldiaceae	<i>Pilea tenuifolia</i> Michx.	<i>Qiu 96128</i> (IND)	<i>Y.-L. Qiu 96128</i> (IND)	EU264279
		Xanthorrhoeaceae	<i>Lomandra obliqua</i> (Thunb.) J. F. Macbr.	<i>Qiu 98016</i> (Z)	<i>Y.-L. Qiu 98016</i> (Z)	EU264240
			<i>Xanthorrhoea quadrangulata</i> F. Muell.	<i>Qiu 97039</i> (IND)	<i>Y.-L. Qiu 97039</i> (IND)	EU264242
Commelinids	Dioscoreales	Dioscoreaceae	<i>Dioscorea</i> L., sp. indet.	97750*A (FTBG)	<i>C. & A. Morton 2009-4</i> (CM)	EU264304
			<i>Tacca chantrieri</i> André	<i>Qiu 01015</i> (MASS)	<i>Y.-L. Qiu 01015</i> (MASS)	EU264278
		Smilacaceae	<i>Smilax china</i> L.	73.0594 (UCBG)	73.0594 (UC, UCBG)	EU264277
		Cyclanthaceae	<i>Carludovica palmata</i> Ruiz. & Pav.	FG4902*A (FTBG)	<i>C. & A. Morton 2009-1</i> (CM)	EU264305
			<i>Talbotia elegans</i> Balf.	61.0758 (UCBG)	61.0758 (UC, UCBG)	EU264280
		Arecaceae	<i>Chamaedorea fragrans</i> Mart.	5875*A (FTBG)	<i>C. & A. Morton 2009-16</i> (CM)	EU264306
		Bromeliaceae	<i>Vriesea splendens</i> (Brongn.) Lem.	<i>Qiu 96073</i> (IND)	<i>Y.-L. Qiu 96073</i> (IND)	EU264239
		Marantaceae	<i>Maranta leuconeura</i> E. Morren	<i>Qiu 95081</i> (IND)	<i>Y.-L. Qiu 95081</i> (IND)	EU264241
		Strelitziaceae	<i>Strelitzia nicolai</i> Regel & K. Koch	77117*A (FTBG)	<i>C. & A. Morton 2009-10</i> (CM)	EU264307
Eudicots	Buxales	Buxaceae	<i>Buxus sempervirens</i> L.	62.0414 (UCBG)	62.0414 (UC, UCBG)	EU264236
			<i>Pachysandra procumbens</i> Michx.	60.1303 (UCBG)	60.1303 (UC, UCBG)	EU264238
	Sabiales	Sabiaceae	<i>Sabia campanulata</i> Wall.	94.0765 (UCBG)	94.0765 (UC, UCBG)	EU264316
	Trochondrales	Trochodendraceae	<i>Trochodendron aralioides</i> Siebold & Zucc.	61.0594 (UCBG)	61.0594 (UC, UCBG)	EU264317

Group	Order	Family	Species	Voucher/Source	Museum voucher/Number	GenBank number
Proteales	Ranunculales	Proteaceae	<i>Platanus orientalis</i> L.	76.0216 (UCBG)	76.0216 (UC, UCBG)	EU264315
		Berberidaceae	<i>Caulophyllum robustum</i> Maxim.	98.0013 (UCBG)	98.0013 (UC, UCBG)	EU264313
Dilleniales	Gunnerales	Eupteleaceae	<i>Nandina domestica</i> Thunb.	no accession listed (PCPT)	<i>C. Morton 2010-4</i> (CM)	EU264309
			<i>Euptelea pleiosperma</i> Hook. f. & Thomson	74.0757 (UCBG)	74.0757 (UC, UCBG)	EU264314
		Lardizabalaceae	<i>Decaisnea fargesii</i> Franch.	2002.0189 (UCBG)	2002.0189 (UC, UCBG)	EU264312
			<i>Lardizabala biternata</i> Ruiz & Pav.	53.116 (UCBG)	53.116 (UC, UCBG)	EU264310
		Papaveraceae	<i>Eschscholzia californica</i> Cham.	62.1419 (UCBG)	62.1419 (UC, UCBG)	EU264311
		Ranunculaceae	<i>Hydrastis canadensis</i> L.	99.059 (UCBG)	99.059 (UC, UCBG)	EU264232
			<i>Nigella damascena</i> L.	87.0734 (UCBG)	87.0734 (UC, UCBG)	EU264308
		Dilleniaceae	<i>Dillenia indica</i> L.	<i>Qiu 95129</i> (IND)	<i>Y.-L. Qiu 95129</i> (IND)	EU264233
		Gunneraceae	<i>Gunnera magellanica</i> Lam.	93.1181 (UCBG)	93.1181 (UC, UCBG)	EU264235
			<i>Gunnera monoica</i> Raoul	<i>Qiu 98071</i> (Z)	<i>Y.-L. Qiu 98071</i> (Z)	EU264234
Caryophyllales		Cactaceae	<i>Pereskia portulacifolia</i> (L.) DC.	8465*A (FTBG)	<i>C. & A. Morton 2009-3</i> (CM)	EU264349
		Caryophyllaceae	<i>Silene nutans</i> L.	<i>M. W. Chase 2292</i> (K)	<i>M. W. Chase 2292</i> (K)	EU264352
			<i>Drosera binata</i> Labill.	71.0223 (UCBG)	71.0223 (UC, UCBG)	EU264347
		Droseraceae	<i>Nepenthes northiana</i> Hook. f.	99.0393 (UCBG)	99.0393 (UC, UCBG)	EU264346
		Nepenthaceae	<i>Bougainvillea glabra</i> Choisy	<i>M. W. Chase 2485</i> (K)	<i>M. W. Chase 2485</i> (K)	EU264351
		Nyctaginaceae	<i>Plumbago auriculata</i> Lam.	87314*A (FTBG)	<i>C. & A. Morton 2009-12</i> (CM)	EU264344
		Plumbaginaceae	<i>Portulaca grandiflora</i> Hook.	no accession listed (PCPT)	<i>C. Morton 2010-5</i> (CM)	EU264350
			<i>Talinum paniculatum</i> (Jacq.) Gaertn.	90.0939 (UCBG)	90.0939 (UC, UCBG)	EU264348
		Portulacaceae	<i>Simmondsia chinensis</i> C. K. Schneid.	<i>Qiu 96120</i> (IND)	<i>Y.-L. Qiu 96120</i> (IND)	EU264345
			<i>Osyris alba</i> L.	2002.0996 (UCBG)	2002.0996 (UC, UCBG)	EU264237
Rosids	Crossosomatales	Simmondsiaceae	<i>Stachyurus praecox</i> Siebold & Zucc.	81.0031 (UCBG)	81.0031 (UC, UCBG)	EU264432
		Santalaceae	<i>Staphylea trifolia</i> L.	512642 (CM)	<i>B. & J. Isaac 18690</i> (CM)	EU264433
		Stachyuraceae	<i>Geranium sanguineum</i> L.	99/1397A (PCPT)	<i>C. Morton 2010-6</i> (CM)	EU264243
		Staphyleaceae	<i>Greyia radlkoferi</i> Szyszyl.	61.1016 (UCBG)	61.1016 (UC, UCBG)	EU264282
		Geraniaceae	<i>Metrosideros nervulosa</i> C. Moore & F. Muell.	<i>M. W. Chase 2451</i> (K)	<i>M. W. Chase 2451</i> (K)	EU264247
		Melianthaceae	<i>Myrciaria cauliflora</i> (Mart.) O. Berg	95/0621 (PCPT)	<i>C. Morton 2010-7</i> (CM)	EU264246
		Myrtaceae	<i>Psidium littorale</i> Raddi	95/0712 (PCPT)	<i>C. Morton 2010-8</i> (CM)	EU264245
		Penaeaceae	<i>Penaea cneorum</i> Meerb.	96.0085 (UCBG)	96.0085 (UC, UCBG)	EU264248
			<i>Liquidambar styraciflua</i> L.	77.0386 (UCBG)	77.0386 (UC, UCBG)	EU264434
		Altingiaceae	<i>Cercidiphyllum japonicum</i> Siebold & Zucc. ex J. J. Hoffm. & J. H. Schult. bis	93.0206 (UCBG)	93.0206 (UC, UCBG)	EU264442
Saxifragales	Cercidiphyllaceae		<i>Daphniphyllum himalayense</i> Müll. Arg.	2002.0188 (UCBG)	2002.0188 (UC, UCBG)	EU264443
			<i>Ribes cynosbati</i> L.	<i>S. P. Grund 2313</i> (CM)	<i>S. P. Grund 2313</i> (CM)	EU264437
			<i>Corylopsis sinensis</i> Hemsl.	81.0149 (UCBG)	81.0149 (UC, UCBG)	EU264436

Appendix 1. Continued.

Group	Order	Family	Species	Voucher/Source	Museum voucher/Number	GenBank number	
Eurosids I	Celastrales	Iteaceae	<i>Hamamelis virginiana</i> L.	<i>L. Speedy LS-07-387</i> (CM)	<i>L. Speedy LS-07-387</i> (CM)	EU264435	
		Paeoniaceae	<i>Itea virginica</i> L.	00/1213 (PCPT)	<i>C. Morton 2010-9</i> (CM)	EU264438	
		Celastraceae	<i>Paeonia lactiflora</i> Pall.	512633 (CM)	<i>B. Isaac 18677</i> (CM)	EU264441	
	Cucurbitales	Celastraceae	<i>Celastrus orbiculatus</i> Thunb.	481520 (CM)	<i>B. & J. Isaac 9632</i> (CM)	EU264428	
			<i>Euonymus atropurpureus</i> Jacq.	512573 (CM)	<i>B. & J. Isaac 18687</i> (CM)	EU264429	
			<i>Parnassia palustris</i> L.	2002.065 (UCBG)	2002.065 (UC, UCBG)	EU264430	
		Begoniaceae	<i>Begonia</i> L., sp. indet.	00/0528 (PCPT)	<i>C. Morton 2010-10</i> (CM)	EU264460	
			Coriariaceae	<i>Coriaria sarmentosa</i> G. Forst.	95.0196 (UCBG)	95.0196 (UC, UCBG)	EU264459
		Cucurbitaceae	<i>Cucumis sativus</i> L.	Morton	<i>C. Morton 2010-29</i> (CM)	EU264458	
			Fabaceae	<i>Acacia</i> Mill., sp. indet.	99/0912 (PCPT)	<i>C. Morton 2010-11</i> (CM)	EU264454
Fabales	Cucurbitales	Fabaceae	<i>Albizia cubana</i> Britton & P. Wilson	2000124*D (FTBG)	<i>C. & A. Morton 2009-21</i> (CM)	EU264455	
			<i>Sophora toromiro</i> Skottsb.	<i>M. W. Chase 747</i> (K)	<i>M. W. Chase 747</i> (K)	EU264456	
			<i>Vicia caroliniana</i> Walter	489294 (CM)	<i>B. & J. Isaac 9682</i> (CM)	EU264457	
	Fagales	Quillajaceae	<i>Quillaja saponaria</i> Molina	36.1419 (UCBG)	36.1419 (UC, UCBG)	EU264453	
		Betulaceae	<i>Betula</i> L., sp. indet.	no accession listed (PCPT)	<i>C. Morton 2010-12</i> (CM)	EU264449	
		Fagaceae	<i>Quercus virginiana</i> Mill.	2000152*A (FTBG)	<i>C. & A. Morton 2009-25</i> (CM)	EU264450	
	Juglandaceae	<i>Juglans guatemalensis</i> W. E. Manning in Standl. & Steyerl.	92.0561 (UCBG)	92.0561 (UC, UCBG)	92.0561 (UC, UCBG)	EU264451	
		Malpighiales	Myricaceae	<i>Platycarya strobilacea</i> Siebold & Zucc.	80.1197 (UCBG)	80.1197 (UC, UCBG)	EU264452
			Chrysobalanaceae	<i>Myrica pensylvanica</i> Hort. Reg. ex Lam.	99/1372 (PCPT)	<i>C. Morton 2010-13</i> (CM)	EU264448
				Clusiaceae	<i>Chrysobalanus icaco</i> L.	99/1920 (PCPT)	<i>C. Morton</i> (no voucher)
Malpighiales	Euphorbiaceae	Clusiaceae	<i>Clusia</i> L., sp. indet.	77.0031 (UCBG)	77.0031 (UC, UCBG)	EU264227	
		Euphorbiaceae	<i>Codiaeum variegatum</i> (L.) Rumph. ex A. Juss.	99/1772 (PCPT)	<i>C. Morton 2010-15</i> (CM)	EU264224	
			<i>Euphorbia nerifolia</i> L.	96/0677 (PCPT)	<i>C. Morton 2010-16</i> (CM)	EU264225	
	Malpighiaceae	<i>Euphorbia pseudocactus</i> A. Berger	96/0678 (PCPT)	<i>C. Morton 2010-17</i> (CM)	EU264226		
		Phyllanthus flexuosus Müll. Arg.	90.0901 (UCBG)	90.0901 (UC, UCBG)	EU264217		
		Byrsonima lucida DC.	69363*A (FTBG)	<i>C. & A. Morton 2009-24</i> (CM)	EU264222		
	Passifloraceae	<i>Malpighia glabra</i> L.	95/0646.A (PCPT)	<i>C. Morton 2010-18</i> (CM)	EU264223		
		<i>Passiflora suberosa</i> L.	<i>S. A. Thompson & J. H. Nishida 2761</i> (CM)	<i>S. A. Thompson & J. H. Nishida 2761</i> (CM)	EU264216		
		Salicaceae	<i>Populus grandidentata</i> Michx.	492481 (CM)	<i>B. Isaac & T. Herman 11781</i> (CM)	EU264220	
			<i>Salix alba</i> L.	512624 (CM)	<i>B. Isaac 18658</i> (CM)	EU264219	
Oxalidales	Violaceae	<i>Hybanthus prunifolius</i> (Humb. & Bonpl.) Schulze-Menz	65657*A (FTBG)	<i>C. & A. Morton 2009-7</i> (CM)	EU264218		
		Cephalotaceae	<i>Cephalotus follicularis</i> Labill.	2001.0116 (UCBG)	2001.0116 (UC, UCBG)	EU264228	
	Cunoniaceae	<i>Cunonia capensis</i> L.	96.0024 (UCBG)	96.0024 (UC, UCBG)	EU264229		

Group	Order	Family	Species	Voucher/Source	Museum voucher/Number	GenBank number
Eurosids II	Rosales	Elaeocarpaceae	<i>Crinodendron patagua</i> Molina	54.1123 (UCBG)	54.1123 (UC, UCBG)	EU264231
		Oxalidaceae	<i>Averrhoa carambola</i> L.	Morton	<i>C. Morton 2010-30</i> (CM)	EU264230
		Cannabaceae	<i>Celtis</i> L., sp. indet.	Morton	<i>C. Morton 2010-31</i> (CM)	EU264265
		Moraceae	<i>Humulus lupulus</i> L.	512631 (CM)	<i>B. Isaac 18675</i> (CM)	EU264283
			<i>Ficus elastica</i> Roxb.	99/1907 (PCPT)	<i>C. Morton 2010-19</i> (CM)	EU264444
			<i>Morus alba</i> L.	512650 (CM)	<i>B. Isaac 18660</i> (CM)	EU264445
			<i>Rhamnus frangula</i> L.	512638 (CM)	<i>B. & J. Isaac 18683</i> (CM)	EU264447
		Rosaceae	<i>Rosa</i> L., sp. indet.	Morton	<i>C. Morton 2010-32</i> (CM)	EU264284
		Ulmaceae	<i>Spiraea japonica</i> L. f.	01/0171 (PCPT)	<i>C. Morton 2010-20</i> (CM)	EU264461
			<i>Zelkova serrata</i> Makino	94.1121 (UCBG)	94.1121 (UC, UCBG)	EU264446
	Malvales	Urticaceae	<i>Urtica dioica</i> L.	99.0492 (UCBG)	99.0492 (UC, UCBG)	EU264281
		Brassicaceae	<i>Arabidopsis thaliana</i> (L.) Heynh.	GenBank NC_003075	GenBank NC_003075	
		Caricaceae	<i>Capparis spinosa</i> L.	62.0728 (UCBG)	62.0728 (UC, UCBG)	EU264420
			<i>Carica papaya</i> L.	volunteer seedlings (FTBG)	<i>C. & A. Morton 2009-26</i> (CM)	EU264421
			<i>Reseda luteola</i> L.	98.0665 (UCBG)	98.0665 (UC, UCBG)	EU264419
		Tropaeolaceae	<i>Tropaeolum majus</i> L.	98.0856 (UCBG)	98.0856 (UC, UCBG)	EU264422
		Bixaceae	<i>Bixa orellana</i> L.	89.0572 (UCBG)	89.0572 (UC, UCBG)	EU264424
		Cistaceae	<i>Cistus parviflorus</i> Lam.	2004.0774 (UCBG)	2004.0774 (UC, UCBG)	EU264426
		Dipterocarpaceae	<i>Helianthemum apenninum</i> Mill.	86.0822 (UCBG)	86.0822 (UC, UCBG)	EU264425
			<i>Hopea odorata</i> Roxb.	961229*A (FTBG)	<i>C. & A. Morton 2009-8</i> (CM)	EU264427
	Sapindales	Malvaceae	<i>Sterculia urens</i> Roxb.	<i>R. Aborn s.n.</i> (CM)	<i>R. Aborn s.n.</i> (CM)	EU264423
		Anacardiaceae	<i>Pistacia lentiscus</i> L.	2002.0618 (UCBG)	2002.0618 (UC, UCBG)	EU264431
			<i>Rhus glabra</i> L.	<i>S. P. Grund & S. V. Parker 3813</i> (CM)	<i>S. P. Grund & S. V. Parker 3813</i> (CM)	EU264400
		Burseraceae	<i>Bursera inaguensis</i> Britton	64269*M (FTBG)	<i>C. & A. Morton 2009-22</i> (CM)	EU264406
	Rutaceae	Meliaceae	<i>Swietenia macrophylla</i> King in Hook.	951002*B (FTBG)	<i>C. & A. Morton 2009-18</i> (CM)	EU264415
		Rutaceae	<i>Helietta parvifolia</i> (A. Gray ex Hemsl.) Benth.	91.0496 (UCBG)	91.0496 (UC, UCBG)	EU264418
			<i>Agathosma ovata</i> (Thunb.) Pillans	<i>M. W. Chase 1667</i> (K)	<i>M. W. Chase 1667</i> (K)	EU264409
			<i>Bosistoa brassii</i> T. G. Hartley	<i>M. W. Chase 2492</i> (K)	<i>M. W. Chase 2492</i> (K)	EU264410
			<i>Casimiroa edulis</i> La Llave	<i>M. W. Chase 1342</i> (K)	<i>M. W. Chase 1342</i> (K)	EU264403
			<i>Citropsis gabunensis</i> Swingle & M. Kellerm.	<i>C. Morton 3286</i> (CRC)	<i>C. Morton 2010-35</i> (CM)	EU264396
			<i>Eremocirus glauca</i> (Lindl.) Swingle	<i>M. W. Chase 1768</i> (K)	<i>M. W. Chase 1768</i> (K)	EU264414
			<i>Fortunella japonica</i> Swingle	<i>M. W. Chase 1758</i> (K)	<i>M. W. Chase 1758</i> (K)	EU264397
			<i>Melicope ternata</i> J. R. Forst. & G. Forst.	<i>M. W. Chase 1759</i> (K)	<i>M. W. Chase 1759</i> (K)	EU264401
			<i>Pilocarpus pennatifolius</i> Lem.	<i>M. W. Chase 1760</i> (K)	<i>M. W. Chase 1760</i> (K)	EU264408
			<i>Poncirus trifoliata</i> (L.) Raf.	<i>C. Morton 27</i> (US)	<i>C. Morton 2010-36</i> (CM)	EU264398

Appendix 1. Continued.

Group	Order	Family	Species	Voucher/Source	Museum voucher/Number	GenBank number
Asterids	Cornales	Sapindaceae	<i>Ravenia infelix</i> Vell.	<i>J. Kallunki et al.</i> 614 (NY)	<i>J. Kallunki et al.</i> 614 (NY)	EU264411
			<i>Ruta graveolens</i> L.	99/1337 (PCPT)	<i>C. Morton</i> 2010-21 (CM)	EU264412
			<i>Spathelia excelsa</i> (K. Krause) R. S. Cowan & Brizicky	<i>M. W. Chase</i> 1899 (K)	<i>M. W. Chase</i> 1899 (K)	EU264413
			<i>Teclea simplicifolia</i> (Engl.) Engl.	<i>M. W. Chase</i> 1764 (K)	<i>M. W. Chase</i> 1764 (K)	EU264417
			<i>Tetradium daniellii</i> (Benn.) T. G. Hartley	<i>M. W. Chase</i> 633 (K)	<i>M. W. Chase</i> 633 (K)	EU264416
			<i>Triphasia trifolia</i> P. Wilson	<i>C. Morton</i> 111G (IPB)	<i>C. Morton</i> 2010-33 (CM)	EU264399
			<i>Zanthoxylum monophyllum</i> P. Wilson	<i>M. W. Chase</i> 332 (K)	<i>M. W. Chase</i> 332 (K)	EU264402
			<i>Acer rubrum</i> L.	512646 (CM)	<i>B. Isaac</i> 18667 (CM)	EU264407
			<i>Aesculus hippocastanum</i> L.	513762 (CM)	<i>B. & J. Isaac</i> 18671 (CM)	EU264404
			<i>Ailanthus alissima</i> (Mill.) Swingle	<i>H. George s.n.</i> (CM)	<i>H. George s.n.</i> (CM)	EU264405
Asterids	Cornales	Simaroubaceae	<i>Cornus mas</i> L.	01/0151 (PCPT)	<i>C. Morton</i> 2010-22 (CM)	EU264354
			<i>Nyssa sylvatica</i> Marshall	82.2016 (UCBG)	82.2016 (UC, UCBG)	EU264353
			<i>Hydrangea paniculata</i> Siebold	512623 (CM)	<i>B. Isaac</i> 18659 (CM)	EU264355
			<i>Clematoclethra lasioclada</i> Maxim.	<i>M. W. Chase</i> 596 (K)	<i>M. W. Chase</i> 596 (K)	EU264325
			<i>Impatiens repens</i> Moon	<i>M. W. Chase</i> 901 (K)	<i>M. W. Chase</i> 901 (K)	EU264366
			<i>Clethra arborea</i> Aiton	<i>M. W. Chase</i> 902 (K)	<i>M. W. Chase</i> 902 (K)	EU264321
			<i>Cyrilla racemiflora</i> L.	80.0255 (UCBG)	80.0255 (UC, UCBG)	EU264320
			<i>Galax urceolata</i> (Poir.) Brummitt	<i>M. W. Chase</i> 1449 (K)	<i>M. W. Chase</i> 1449 (K)	EU264335
			<i>Diospyros digyna</i> Jacq.	X.3-28*A (FTBG)	<i>C. & A. Morton</i> 2009-9 (CM)	EU264323
			<i>Diospyros mespiliformis</i> Hochst. ex A. DC.	6190*A (FTBG)	<i>C. & A. Morton</i> 2009-20 (CM)	EU264322
Asterids	Cornales	Ebenaceae	<i>Euclea natalensis</i> A. DC.	<i>P. Goldblatt</i> 9275 (MO)	<i>P. Goldblatt</i> 9275 (MO)	EU264324
			<i>Craibiodendron yunnanense</i> W. W. Sm.	<i>M. W. Chase</i> 908 (K)	<i>M. W. Chase</i> 908 (K)	EU264343
			<i>Rhododendron hippophaeoides</i> Balf. f. & W. W. Sm.	RBG Edinburgh 321022 (E)	<i>Morton</i> (no voucher)	EU264319
			<i>Fouquieria columnaris</i> Kellogg ex Curran	<i>M. W. Chase</i> 269 (NCU)	<i>M. W. Chase</i> 269 (NCU)	EU264359
			<i>Fouquieria splendens</i> Engelm. in Wisl.	2000202*A (FTBG)	<i>C. & A. Morton</i> 2009-5 (CM)	EU264361
			<i>Couropita guianensis</i> Aubl.	X.1-489*A, Plot 137 (FTBG)	<i>C. & A. Morton</i> 2009-14 (CM)	EU264332
			<i>Marcgravia rectiflora</i> Triana & Planch.	<i>M. W. Chase</i> 331 (NCU)	<i>M. W. Chase</i> 331 (NCU)	EU264329
			<i>Marcgravia rectiflora</i> Triana & Planch.	941185*A (FTBG)	<i>Morton</i> (no voucher)	EU264362
			<i>Ternstroemia impressa</i> Lundell	<i>P. Fritsch</i> 282	<i>P. W. Fritsch</i> 1771 (CAS)	EU264339
			<i>Ternstroemia stahliae</i> Krug & Urb.	<i>F. Axelrod</i> 4538 (UPR)	<i>F. Axelrod</i> 4538 (UPR)	EU264357
Asterids	Cornales	Primulaceae	<i>Androsace spinulifera</i> R. Knuth	<i>M. W. Chase</i> 954 (NCU)	<i>M. W. Chase</i> 954 (NCU)	EU264327
			<i>Clavija eggersiana</i> Mez	<i>M. W. Chase</i> 216 (NCU)	<i>M. W. Chase</i> 216 (NCU)	EU264338
			<i>Clavija euerganea</i> J. F. Macbr.	86197*A (FTBG)	<i>C. & A. Morton</i> 2009-19 (CM)	EU264363
			<i>Elingamita johnsonii</i> G. T. S. Baylis	<i>M. W. Chase</i> 958 (K)	<i>M. W. Chase</i> 958 (K)	EU264358
			<i>Maesa myrsinoides</i> H. Lév.	<i>M. W. Chase</i> 309 (NCU)	<i>M. W. Chase</i> 309 (NCU)	EU264330

Group	Order	Family	Species	Voucher/Source	Museum voucher/Number	GenBank number
Euasterids I	Sapotales	Sapotaceae	<i>Parathesis</i> (A. DC.) Hook. f., sp. indet.	<i>P. Fritsch</i> 284	<i>P. W. Fritsch</i> 1770 (CAS)	EU264365
			<i>Manilkara zapota</i> (L.) P. Royen	PRC137*B (FTBG)	<i>C. & A. Morton</i> 2009-11 (CM)	EU264333
			<i>Planchonella novozelandica</i> (F. Muell.) Allan	<i>P. Fritsch</i> 287	<i>P. W. Fritsch</i> 1769 (CAS)	EU264331
	Sarraceniales	Sarraceniaceae	<i>Heliamphora heterodoxa</i> Steyerl.	71.0222 (UCBG)	71.0222 (UC, UCBG)	EU264342
			<i>Sarracenia flava</i> L.	<i>M. W. Chase</i> 144 (NCB)	<i>M. W. Chase</i> 144 (NCB)	EU264341
			<i>Rehderodendron macrocarpum</i> Hu	<i>M. W. Chase</i> 911 (K)	<i>M. W. Chase</i> 911 (K)	EU264340
	Styracinales	Styracaceae	<i>Styrax japonicus</i> Siebold & Zucc.	<i>M. W. Chase</i> 974 (K)	<i>M. W. Chase</i> 974 (K)	EU264336
			<i>Symplocos costata</i> Choisy ex Zoll.	<i>M. W. Chase</i> 1374 (K)	<i>M. W. Chase</i> 1374 (K)	EU264337
			<i>Symplocos paniculata</i> Miq.	<i>M. W. Chase</i> 913 (K)	<i>M. W. Chase</i> 913 (K)	EU264360
	Tetrameristales	Tetrameristaceae	<i>Pelliciera rhizophorae</i> Planch. & Triana	<i>R. T. Pennington et al.</i> 586	<i>R. T. Pennington et al.</i> 586 (E)	EU264318
			<i>Tetramerista</i> Miq., sp. indet.	<i>M. Coode</i> 7925 (K)	<i>M. Coode</i> 7925 (K)	EU264326
			<i>Cleyera japonica</i> Thunb.	<i>M. W. Chase</i> 1690 (K)	<i>M. W. Chase</i> 1690 (K)	EU264364
	Theaceales	Theaceae	<i>Eurya japonica</i> Thunb.	<i>M. W. Chase</i> 1448 (K)	<i>M. W. Chase</i> 1448 (K)	EU264328
			<i>Schima superba</i> Gardner & Champ.	<i>Lammers s.n.</i> (F)	<i>Lammers s.n.</i> (F)	EU264356
			<i>Stewartia pseudocamellia</i> Maxim.	<i>M. W. Chase</i> 984 (K)	<i>M. W. Chase</i> 984 (K)	EU264334
	Boraginiales	Boraginaceae	<i>Borago officinalis</i> L.	73.0653 (UCBG)	73.0653 (UC, UCBG)	EU264371
			<i>Eucommia ulmoides</i> Oliv.	86.128 (UCBG)	86.128 (UC, UCBG)	EU264381
			<i>Garrya elliptica</i> Douglas ex Lindl.	64.1209 (UCBG)	64.1209 (UC, UCBG)	EU264382
	Gentianales	Apocynaceae	<i>Nerium oleander</i> L.	78779*B (FTBG)	<i>C. & A. Morton</i> 2009-23 (CM)	EU264377
			<i>Gelsemium sempervirens</i> (L.) J. St.-Hil.	92.032 (UCBG)	92.032 (UC, UCBG)	EU264383
			<i>Strychnos nux-vomica</i> L.	<i>M. W. Chase</i> 2538	<i>M. W. Chase</i> 2538	EU264369
	Lamiales	Bignoniaceae	<i>Coffea arabica</i> L.	99/0283 (PCPT)	<i>C. Morton</i> 2010-23 (CM)	EU264367
			<i>Luculia gratissima</i> Wall.	54.1342 (UCBG)	54.1342 (UC, UCBG)	EU264384
			<i>Catalpa bignonioides</i> Walter	512590 (CM)	<i>B. & J. Isaac</i> 18672 (CM)	EU264375
	Gesneriales	Calceolariaceae	<i>Calceolaria tomentosa</i> Ruiz & Pav.	36.1783 (UCBG)	36.1783 (UC, UCBG)	EU264286
			<i>Saintpaulia grotei</i> Engl.	53.0387 (UCBG)	53.0387 (UC, UCBG)	EU264376
			<i>Antirrhinum</i> L., sp. indet.	no accession listed (PCPT)	<i>C. Morton</i> 2010-24 (CM)	EU264368
	Solanales	Convolvulaceae	<i>Erinus alpinus</i> L.	00/0568 (PCPT)	<i>C. Morton</i> 2010-25 (CM)	EU264288
			<i>Ipomoea batatas</i> (L.) Lam.	88.056 (UCBG)	88.056 (UC, UCBG)	EU264386
			<i>Montinia caryophyllacea</i> Thunb.	75.0110 (UCBG)	75.0110 (UC, UCBG)	EU264379
Euasterids II	Bruniales	Solanaceae	<i>Solanum carolinense</i> L.	380780 (CM)	<i>B. Isaac & C. Chuey</i> 4574 (CM)	EU264370
			<i>Berzelia lanuginosa</i> (L.) Brongn.	96.0009 (UCBG)	96.0009 (UC, UCBG)	EU264380
			<i>Escallonia viscosa</i> Forbes	54.0058 (UCBG)	54.0058 (UC, UCBG)	EU264394
	Apiales	Apiaceae	<i>Apium graveolens</i> L.	<i>B. Isaac</i> 18674 (CM)	<i>B. Isaac</i> 18674 (CM)	EU264378
			<i>Heteromorpha arborescens</i> Cham. & Schltdl.	76.0712 (UCBG)	76.0712 (UC, UCBG)	EU264385
			<i>Hydrocotyle verticillata</i> Thunb.	82.1355 (UCBG)	82.1355 (UC, UCBG)	EU264291

Appendix 1. Continued.

Group	Order	Family	Species	Voucher/Source	Museum voucher/Number	GenBank number
Aquifoliales	Araliaceae		<i>Hedera</i> L., sp. indet.	Morton	<i>C. Morton 2010-34</i> (CM)	EU264372
			<i>Panax quinquefolius</i> L.	86.1138 (UCBG)	86.1138 (UC, UCBG)	EU264289
			<i>Pittosporum daphniphylloides</i> Hayata	75.0918 (UCBG)	75.0918 (UC, UCBG)	EU264290
			<i>Ilex glabra</i> A. Gray	99/1498 (PCPT)	<i>C. Morton 2010-26</i> (CM)	EU264373
Asterales	Helwingiaceae		<i>Helwingia japonica</i> (Thunb.) F. Dietr.	2001.0229 (UCBG)	2001.0229 (UC, UCBG)	EU264374
			<i>Centauria montana</i> L.	87.0398 (UCBG)	87.0398 (UC, UCBG)	EU264395
Dipsacales	Campanulaceae		<i>Campanula rapunculus</i> L.	99.0118 (UCBG)	99.0118 (UC, UCBG)	EU264393
			<i>Stylidium adnatum</i> R. Br.	71.0519 (UCBG)	71.0519 (UC, UCBG)	EU264392
			<i>Viburnum prunifolium</i> L.	512643 (CM)	<i>B. & J. Isaac 18691</i> (CM)	EU264390
			<i>Dipelta yunnanensis</i> Franch.	99.045 (UCBG)	99.045 (UC, UCBG)	EU264387
	Caprifoliaceae		<i>Lonicera japonica</i> Thunb. ex Murray	BED427 (UCBG)	BED427 (UC, UCBG)	EU264389
			<i>Patrinia scabiosifolia</i> Link	93.1081 (UCBG)	93.1081 (UC, UCBG)	EU264391

Abbreviations used for the following botanical gardens: CM, Carnegie Museum of Natural History Herbarium; CRC, Citrus Genetic Resources, Riverside; FTBG, Fairchild Tropical Botanic Garden; IPB, Bogor Botanical Garden; PCPT, Phipps Conservatory and Botanical Gardens; UCBG, University of California Botanical Garden at Berkeley. Other abbreviations follow Index Herbariorum.